JEAN VOEDTS

COURS DE MATHÉMATIQUES

MP-MP*

Cours de Mathématiques

MP-MP*

Jean VOEDTS
Agrégé de mathématiques
Professeur en mathématiques spéciales
au lycée Faidherbe, Lille



PRÉFACE

Cet ouvrage est issu d'un cours polycopié utilisé depuis quelques années par les élèves de la classe de Spéciales MP* du Lycée FAIDHERBE de LILLE. Il suit en conséquence d'assez près le programme de Mathématiques des classes préparatoires MP/MP* et est aussi destiné aux étudiants de premier cycle universitaire.

Il peut être utilisé comme outil de référence, puisqu'il contient toutes les définitions, les théorèmes du cours et leurs démonstrations. A ce titre, il pourra également être utile aux candidats au CAPES ou à l'Agrégation.

De nombreux exercices, illustrant les notions de base du cours, sont aussi traités à titre d'exemple. En outre, beaucoup de compléments sont donnés sous forme d'exercices progressifs (par exemple, ce qui touche aux transformations de Fourier et Laplace, aux fonctions d'une variable complexe) pour ne pas alourdir le cours mais donner cependant un élargissement vers des notions importantes utilisant directement les théorèmes du programme. Il y a ainsi, dans cet ouvrage, un peu plus que ce qui peut être enseigné en un an à un étudiant de classe préparatoire ou de DEUG. A la fin de chaque chapitre, on trouvera également une liste d'exercices tout à fait abordables lorsque le cours est connu.

Je tiens à remercier particulièrement mon collègue et ami Philippe Royer pour le soin qu'il a apporté à la lecture du manuscrit et pour ses nombreuses remarques qui m'ont permis d'améliorer le contenu de cet ouvrage. Mes pensées vont également vers les élèves de la classe de MP* où j'ai la chance d'enseigner. Grâce à leur lecture attentive du polycopié, de nombreuses erreurs ont pu être rectifiées. Qu'ils en soient ici sincèrement remerciés. Enfin, je n'oublierai pas mes filles Nadia et Sophia, qui ont sans doute trop vu leur papa s'initier, devant l'ordinateur, aux merveilles de LaTeX. Un grand, grand merci pour leur compréhension et la patience de Najia, leur maman.

Jean VOEDTS

TABLE DES MATIÈRES

Chapitre 1

Espaces vectoriels

1-1 Structure d'espace vectoriel

Dans tout ce qui suit, $\mathbb{K}$ désigne un corps commutatif quelconque $\mathbb{R}$. Dans la pratique ce sera un sous-corps de $\mathbb{C}$ (et très souvent $\mathbb{C}$, $\mathbb{R}$, ou $\mathbb{Q}$). 0 est son élément neutre pour l'addition, 1 l'est pour la multiplication.

1-1.1 Notion de $\mathbb{K}$-espace vectoriel

Soit $\mathbb{E}$ un ensemble non vide, dont les éléments seront appelés vecteurs (par opposition aux éléments de $\mathbb{K}$ qu'on appelle traditionnellement les scalaires). Par commodité, les vecteurs seront notés (en général) avec des lettres de l'alphabet latin, les scalaires étant représentés par des lettres grecques.

DÉFINITION 1-1.1 *Une structure de $\mathbb{K}$-ev sur $\mathbb{E}$ est déterminée par :*

· *Une loi de composition interne (donc une application $\mathbb{E} \times \mathbb{E} \to \mathbb{E}$) notée + telle que $(\mathbb{E}, +)$ soit un groupe commutatif (dont l'élément neutre est le vecteur nul noté $\mathbf{0}_{\mathbb{E}}$).*

· *Une loi de composition externe à domaine d'opérateurs égal à $\mathbb{K}$, c'est-à-dire une application $\mathbb{K} \times \mathbb{E} \to \mathbb{E}$, $(\lambda, x) \mapsto \lambda x$, appelée multiplication externe, vérifiant les propriétés :*

$$\begin{array}{ll} \forall \alpha, \beta \in \mathbb{K} \quad \forall x \in \mathbb{E} & (\alpha + \beta) x = \alpha x + \beta x \\ \forall \alpha \in \mathbb{K} \quad \forall x, y \in \mathbb{E} & \alpha (x + y) = \alpha x + \alpha y \\ \forall x \in \mathbb{E} & 1x = x \\ \forall \alpha, \beta \in \mathbb{K} \quad \forall x \in \mathbb{E} & \alpha (\beta x) = (\alpha\beta) x \end{array}$$

EXEMPLE 1-1.2 Si $\mathbb{L}$ est un corps commutatif dont $\mathbb{K}$ est un sous-corps, $\mathbb{L}$ est muni d'une structure naturelle de $\mathbb{K}$-ev, pour laquelle la multiplication externe n'est autre que la restriction à $\mathbb{K} \times \mathbb{L}$ de la multiplication interne dans $\mathbb{L}$. Ainsi $\mathbb{C}$ peut être considéré comme un $\mathbb{R}$-ev et comme un $\mathbb{Q}$-ev. Il s'agit là de deux structures totalement différentes.

EXEMPLE 1-1.3 Si $\mathbb{L}$ et $\mathbb{K}$ sont comme dans l'exemple précédent, tout $\mathbb{L}$-ev peut être muni d'une structure de $\mathbb{K}$-ev, en restreignant le corps des scalaires. Il est donc nécessaire, lorsqu'une confusion est possible, de bien préciser le corps des scalaires.

EXEMPLE 1-1.4 Si $\mathbb{K}$ est un corps commutatif, l'anneau $\mathbb{K}[X]$ des polynômes à une indéterminée à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-ev.

EXEMPLE 1-1.5 Si $\mathbb{X}$ est un ensemble non vide, l'ensemble $\mathbb{K}^{\mathbb{X}}$ des fonctions de $\mathbb{X}$ dans $\mathbb{K}$ est un $\mathbb{K}$-ev pour les lois naturelles

$$\forall f, g \in \mathbb{K}^{\mathbb{X}}\ \forall \lambda \in \mathbb{K}\ \forall x \in \mathbb{X} \quad (f+g)(x) = f(x) + g(x) \quad \text{et} \quad (\lambda f)(x) = \lambda f(x)$$

Plus généralement, si $\mathbb{E}$ est un $\mathbb{K}$-ev, $\mathbb{E}^{\mathbb{X}}$ possède également une structure naturelle d'espace vectoriel.

EXEMPLE 1-1.6 En particulier, pour $n \in \mathbb{N}^*$, $\mathbb{K}^n$ est un $\mathbb{K}$-ev pour les opérations

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$
$$\lambda (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$$

EXERCICE 1-1.7 En développant de deux manières différentes $(1+1)(x+y)$ montrer que la commutativité de l'addition est conséquence des autres axiomes de définition de la structure d'espace vectoriel.

1-1.2 Règles de calcul dans un espace vectoriel

En développant $(1+0)\,x$, $\lambda\,(x + 0_{\mathbb{E}})$, $(\lambda + (-\lambda))\,x$ et $\lambda\,(x + (-x))$, et en utilisant les règles de simplification dans le groupe additif $(\mathbb{E}, +)$, on obtient

$$0.x = \lambda.0_{\mathbb{E}} = 0_{\mathbb{E}} \qquad (*)$$

$$\lambda(-x) = (-\lambda)\,x = -\,(\lambda x)$$

La propriété $(*)$ possède la réciproque suivante

PROPOSITION 1-1.8 Si $\alpha \in \mathbb{K}$ et $x \in \mathbb{E}$ $\quad \alpha x = 0_{\mathbb{E}} \Longrightarrow \alpha = 0$ **ou** $x = 0_{\mathbb{E}}$

Démonstration : Si $\alpha \neq 0$ il suffit de calculer

$$\alpha^{-1}(\alpha x) = (\alpha^{-1}\alpha)\,x = x = 0_{\mathbb{E}} \quad \blacksquare$$

REMARQUE 1-1.9 Le fait que $\mathbb{K}$ soit un corps est ici utilisé pour être assuré de l'inversibilité (pour la multiplication) de tout scalaire non nul. La notion d'espace vectoriel peut être généralisée en remplaçant le corps $\mathbb{K}$ par un anneau commutatif $\mathbb{A}$. On parle alors de $\mathbb{A}$-module. Dans ce cas, la proposition (**1-1**.8) n'est plus valable en général. Il en résulte que les notions de dépendance et d'indépendance linéaires sont plus délicates à manipuler dans le cadre des modules. Par exemple, un groupe commutatif $(\mathbb{G}, +)$ peut être muni d'une structure de $\mathbb{Z}$-module, en posant comme d'habitude

$$\text{Pour } n \in \mathbb{Z} \text{ et } g \in \mathbb{G} \qquad n.g = \begin{cases} g + \cdots + g \quad (n \text{ fois}) \quad \text{si } n > 0 \\ (-g) + \cdots + (-g) \quad (-n \text{ fois}) \quad \text{si } n < 0 \end{cases}$$

On sait alors que $ng = 0_{\mathbb{G}}$ signifie que n est un multiple de l'ordre de g dans le groupe[1] $\mathbb{G}$, et n'entraîne pas nécessairement $n = 0$ si $g \neq 0_{\mathbb{G}}$.

EXERCICE 1-1.10 Que peut-on déduire de l'égalité $\alpha x = \beta x$? de $\alpha x = \alpha y$?

1-1.3 Combinaison linéaire d'une famille de vecteurs

1-1.3.1 Cas d'une famille finie

Si I est un ensemble (d'indices) fini, on appelle **famille de vecteurs de $\mathbb{E}$ indexée par I** toute application $I \to \mathbb{E}$ définie par $i \mapsto x_i$, qu'on notera

$$\mathcal{F} = (x_i)_{i \in I}$$

On appelle cardinal de cette famille le cardinal de l'ensemble I. Ce cardinal peut être strictement supérieur au cardinal de la partie $\{x_i \mid i \in I\}$, si l'application $i \mapsto x_i$ est non injective. Lorsque I est de cardinal n, l'existence d'une bijection entre I et $\{1, \ldots, n\}$ fait qu'on utilisera le plus souvent $\{1, \ldots, n\}$ comme ensemble d'indices.

Si $(x_i)_{1 \leqslant i \leqslant n}$ est une famille de vecteurs de $\mathbb{E}$ (n est un entier naturel non nul), un vecteur $x \in \mathbb{E}$ est dit **combinaison linéaire de la famille $(x_i)_{1 \leqslant i \leqslant n}$** s'il existe une famille de scalaires $(\alpha_i)_{1 \leqslant i \leqslant n}$ tels que

$$x = \sum_{i=1}^{n} \alpha_i x_i$$

Les règles de calcul dans un espace vectoriel montrent immédiatement le résultat suivant :

PROPOSITION 1-1.11 Si p vecteurs $y_1, \ldots, y_p$ de $\mathbb{E}$ sont combinaisons linéaires d'une famille $(x_i)_{1 \leqslant i \leqslant n}$, toute combinaison linéaire de $(y_i)_{1 \leqslant i \leqslant p}$ est combinaison linéaire de $(x_i)_{1 \leqslant i \leqslant n}$.

[1] Voir section **17-1.3.3**.

1-1.3.2 Cas d'une famille infinie

Comme précédemment, si I est un ensemble (d'indices) infini, on appelle **famille de vecteurs de $\mathbb{E}$ indexée par I** toute application $I \to \mathbb{E}$ définie par $i \mapsto x_i$, qu'on notera $\mathcal{F} = (x_i)_{i \in I}$. Il ne faut pas confondre cette famille avec l'image de l'application considérée $\{x_i,\ i \in I\}$. En particulier l'application $i \mapsto x_i$ n'est pas injective en général.

DÉFINITION 1-1.12 *Un vecteur $x \in \mathbb{E}$ est dit combinaison linéaire de la famille $\mathcal{F} = (x_i)_{i \in I}$ si et seulement si il existe une partie* ***finie*** *$J = \{i_1, \ldots, i_p\} \subset I$ telle que x soit combinaison linéaire de la famille finie $(x_{i_k})_{1 \leqslant k \leqslant p}$. Une combinaison linéaire de $\mathcal{F}$ ne fait donc intervenir qu'un nombre fini de vecteurs de $\mathcal{F}$, elle est combinaison linéaire d'une sous-famille* ***finie*** *extraite de $\mathcal{F}$.*

EXEMPLE 1-1.13 Dans l'espace $\mathbb{K}[X]$, tout polynôme est combinaison linéaire de la famille $(X^i)_{i \in \mathbb{N}}$, mais un polynôme donné ne fait intervenir qu'un nombre fini de monômes.

Par abus de notation, si x s'écrit comme précédemment

$$x = \alpha_{i_1} x_{i_1} + \cdots + \alpha_{i_p} x_{i_p}$$

on notera une telle combinaison linéaire

$$x = \sum_{i \in I} \alpha_i x_i$$

en considérant que les scalaires $(\alpha_i)_{i \in I}$ sont tous nuls, sauf un nombre fini (on dira que **ces scalaires sont presque tous nuls**), et en ne tenant compte dans le calcul que des termes pour lesquels $\alpha_i \neq 0$. C'est assez cohérent puisqu'on ne tient pas compte de $0_{\mathbb{E}}$, élément neutre pour l'addition dans $\mathbb{E}$, mais il s'agit quand même d'un abus (commode) d'écriture puisque les règles du jeu dans $(\mathbb{E}, +)$ ne permettent d'additionner qu'un nombre fini de vecteurs : 2, 10, $6\ 10^{23}$ si on veut, mais pas une infinité[2] ! Bien sûr, si **tous** les α_i sont nuls, on posera $x = 0_{\mathbb{E}}$.

Pour reprendre l'exemple **1-1**.13, tout polynôme de $\mathbb{K}[\mathbb{X}]$ peut s'écrire

$$P = \sum_{n \in \mathbb{N}} \alpha_n X^n \quad (\text{qu'on note aussi } \sum_{n=0}^{+\infty} \alpha_n X^n)$$

Pour un polynôme P donné, la suite des coefficients s'annule à partir de l'entier $n = d^{\circ}(P) + 1$.

Avec l'abus d'écriture précédent, un peu de réflexion montre qu'on peut calculer sur les combinaisons linéaires comme on le fait dans le cas des familles finies.

[2] En analyse, lorsqu'on considérera une série convergente (de nombres réels par exemple) de terme général u_n, on notera sa somme $\sum_{n=0}^{+\infty} u_n$, mais il s'agit d'une notation qui représente, à proprement parler, non pas une somme d'une infinité de réels, mais la limite d'une suite convergente (celle des sommes partielles de la série). Le contexte n'est pas purement algébrique, mais est lié à la topologie de $\mathbb{R}$. On verra d'ailleurs que l'ordre d'apparition des différents termes peut avoir son importance !

Si $(\alpha_i)_{i\in I}$ et $(\beta_i)_{i\in I}$ sont des familles de scalaires presque tous nuls, il en est de même de $(\alpha_i+\beta_i)_{i\in I}$ et on a

$$\sum_{i\in I}\alpha_i x_i+\sum_{i\in I}\beta_i x_i=\sum_{i\in I}(\alpha_i+\beta_i)\,x_i$$

On a de même

$$\lambda\left(\sum_{i\in I}\alpha_i x_i\right)=\sum_{i\in I}(\lambda\alpha_i)\,x_i$$

C'est bien ainsi par exemple que l'on mène les calculs dans $\mathbb{K}[\mathbb{X}]$, en travaillant avec la famille $(X^n)_{n\in\mathbb{N}}$.

1-1.4 Produit d'un nombre fini d'espaces vectoriels

Si $(\mathbb{E}_i)_{1\leqslant i\leqslant n}$ est une famille finie d'espaces vectoriels **sur le même corps commutatif** $\mathbb{K}$, le produit cartésien $\mathbb{E}_1\times\cdots\times\mathbb{E}_n$ peut être muni d'une structure naturelle de $\mathbb{K}$−ev, les calculs étant menés "coordonnée par coordonnée" :

$$(x_1,\ldots,x_n)+(y_1,\ldots,y_n)=(x_1+y_1,\ldots,x_n+y_n)$$
$$\lambda\,(x_1,\ldots,x_n)=(\lambda x_1,\ldots,\lambda x_n)$$

REMARQUE 1-1.14 Si $(\mathbb{E}_i)_{i\in I}$ est une famille infinie de $\mathbb{K}$-ev, le produit $\prod_{i\in I}\mathbb{E}_i$ est défini comme étant l'ensemble des applications $i\mapsto x_i$ de I dans $\bigcup_{i\in I}\mathbb{E}_i$, avec

$$\forall i\in I\quad x_i\in E_i$$

Une telle application sera notée $(x_i)_{i\in I}$. Ici encore, en raisonnant coordonnée par coordonnée, on peut munir $\prod_{i\in I}\mathbb{E}_i$ d'une structure de $\mathbb{K}$-ev. Les espaces $\mathbb{K}^n$ et $\mathbb{K}^{\mathbb{X}}$ donnés en exemple au début de cette section sont des produits de ce type (respectivement $\prod_{1\leqslant i\leqslant n}\mathbb{K}$ et $\prod_{x\in\mathbb{X}}\mathbb{K}$).

1-2 Sous-espace vectoriel

1-2.1 Définition

DÉFINITION 1-2.1 *Une partie $\mathbb{F}$ d'un espace vectoriel $\mathbb{E}$ est dite sous-espace vectoriel de $\mathbb{E}$ si et seulement si $\mathbb{F}$ est non vide et est stable pour les deux lois de l'espace $\mathbb{E}$, c'est-à-dire*

$$\forall x,y\in\mathbb{F}\quad x+y\in F\quad \text{et}\quad \forall\lambda\in\mathbb{K}\ \forall x\in\mathbb{F}\quad \lambda.x\in\mathbb{F}$$

Il résulte alors immédiatement de la définition précédente que $(\mathbb{F}, +, \cdot)$ est lui-même un $\mathbb{K}$-ev. En effet, on a stabilité de $\mathbb{F}$ pour l'addition et passage à l'opposé puisque $x \in \mathbb{F} \Rightarrow (-1).x = -x \in \mathbb{F}$. $(\mathbb{F}, +)$ est donc un sous-groupe de $(\mathbb{E}, +)$, et les propriétés de la loi externe, valables sur $\mathbb{K} \times \mathbf{E}$ sont valables en restriction sur $\mathbb{K} \times \mathbb{F}$.

Pour montrer qu'une partie $\mathbb{F}$ de $\mathbb{E}$ est un sous-espace vectoriel, il faut d'abord montrer que $\mathbb{F} \neq \emptyset$. Ceci se fait en général en montrant que $0_{\mathbf{E}} \in \mathbb{F}$, puisque l'élément neutre de $(\mathbb{E}, +)$ est aussi celui de $(\mathbb{F}, +)$. On remarquera que $\{0_{\mathbf{E}}\}$ et $\mathbb{E}$ sont toujours des sous-espaces de $\mathbb{E}$. Les autres sous-espaces éventuels de $\mathbf{E}$ seront dits non-triviaux.

Plutôt que de séparer les lois interne et externe, on utilise souvent la caractérisation suivante :

PROPOSITION 1-2.2 Une partie non vide $\mathbb{F}$ d'un espace vectoriel est un sous-espace vectoriel si et seulement si

$$\forall \alpha, \beta \in \mathbb{K} \quad \forall x, y \in \mathbb{F} \qquad \alpha x + \beta y \in \mathbb{F}$$

(ou, de manière équivalente, $x + \alpha y \in \mathbb{F}$)

Par associativité de l'addition interne, et comme les combinaisons linéaires ne font toujours intervenir qu'un nombre fini de vecteurs, on en déduit le

COROLLAIRE 1-2.3 Une partie non vide $\mathbb{F}$ d'un espace vectoriel est un sous-espace vectoriel si et seulement si les combinaisons linéaires de toute famille de vecteurs de $\mathbb{F}$ appartiennent à $\mathbb{F}$.

Un sous-espace vectoriel est donc une partie "fermée" pour les opérations de combinaison linéaire.

On se souviendra que, pour montrer qu'un ensemble muni d'une loi interne et d'une loi externe à domaine d'opérateurs dans un corps commutatif $\mathbb{K}$ est un espace vectoriel, il est souvent commode de montrer que c'est un sous-espace vectoriel d'une structure plus vaste **dont il est "bien connu" qu'elle est un espace vectoriel.** Cela évite des vérifications lourdes et inutiles sur les lois de composition.

1-2.2 Intersection de s.e.v. Sous-espace engendré par une famille

Le théorème suivant résulte immédiatement de la caractérisation des s.e.v. :

THÉORÈME 1-2.4 Si $(\mathbf{F}_i)_{i \in I}$ est une famille (non vide) de sous-espaces vectoriels de $\mathbb{E}$, $\bigcap_{i \in I} \mathbf{F}_i$ est un sous-espace vectoriel de $\mathbb{E}$.

Ce théorème est à la base de la notion de s.e.v engendré par une famille de vecteurs (il en sera de même lorsque, dans un groupe, il s'agira de définir le sous-groupe engendré par une famille ; on verra ultérieurement dans le cours la notion d'idéal engendré par une famille, dans un anneau commutatif).

THÉORÈME 1-2.5 (et définition) Si $\mathcal{F} = (x_i)_{i \in I}$ est une famille de vecteurs d'un espace vectoriel $\mathbb{E}$, il existe un plus petit (au sens de l'inclusion) sous-espace vectoriel de $\mathbb{E}$ contenant tous les vecteurs de $\mathcal{F}$. On l'appelle sous-espace vectoriel engendré par $\mathcal{F}$ et on le note vect($\mathcal{F}$).

Démonstration : La famille $\mathbb{X}$ des sous-espaces de $\mathbb{E}$ contenant $\mathcal{F}$ est non vide (puisqu'elle contient $\mathbb{E}$). On considère

$$\mathbb{F} = \bigcap_{\mathbb{G} \in \mathbb{X}} \mathbb{G}$$

C'est un sous-espace de $\mathbb{E}$ d'après le théorème (**1-2**.4). Par construction, il contient tous les vecteurs de $\mathcal{F}$. Etant inclus dans tous les éléments de $\mathbb{X}$, il est évidemment le plus petit (au sens de l'inclusion) possédant cette propriété. ■

Il s'agit là d'une construction "par l'extérieur" de vect($\mathcal{F}$). La notion de combinaison linéaire permet de caractériser les éléments de vect($\mathcal{F}$), et donc d'en donner une construction "par l'intérieur" :

THÉORÈME 1-2.6 Si $\mathcal{F} = (x_i)_{i \in I}$ est une famille non vide de vecteurs d'un espace vectoriel $\mathbb{E}$, vect($\mathcal{F}$) est l'ensemble des combinaisons linéaires des vecteurs de $\mathcal{F}$.

Démonstration : On vérifie aisément que

$$\mathbb{A} = \{x \in \mathbb{E} \mid x \text{ est combinaison linéaire de } \mathcal{F}\}$$

est un sous-espace de $\mathbb{E}$ contenant $\mathcal{F}$. Comme tout sous-espace est stable par combinaisons linéaires, tout sous-espace de $\mathbb{E}$ contenant $\mathcal{F}$ contient $\mathbb{A}$, qui est donc bien le plus petit (au sens de l'inclusion) possédant cette propriété. ■

REMARQUE 1-2.7 Un ensemble $\mathbb{X}$ peut toujours être considéré, par abus de langage, comme une famille indexée par lui-même :

$$\mathbb{X} \simeq (x)_{x \in \mathbb{X}}$$

(l'indexation est ici injective). C'est d'ailleurs cette identification qui a été implicite dans la démonstration du théorème (**1-2**.5). On pourra donc parler également de combinaisons linéaires de vecteurs d'une partie de $\mathbb{E}$, et du s.e.v. engendré par une partie de $\mathbb{E}$. En particulier, le plus petit s.e.v. de $\mathbb{E}$ étant $\{0_{\mathbb{E}}\}$, on a

$$\text{vect}(\emptyset) = \{0_{\mathbb{E}}\}$$

De même

$$A \subset B \Rightarrow \text{vect}(A) \subset \text{vect}(B)$$

EXERCICE 1-2.8 Montrer que la réunion de deux s.e.v. n'est un sous-espace vectoriel que si l'un des deux sous-espaces contient l'autre.

EXERCICE 1-2.9 Si $\mathbb{K}$ est un corps infini, montrer qu'une union finie de sous-espaces n'est un sous-espace que si l'un des s.e.v. contient tous les autres. (*Indication* : raisonner par récurrence sur le nombre de sous-espaces. Si la propriété est démontrée au rang $n-1$, et si on suppose que $\mathbb{F}_1 \cup \cdots \cup \mathbb{F}_n$ est un s.e.v., considérer des vecteurs $x \in \mathbb{F}_n - \mathbb{F}_1 \cup \cdots \cup \mathbb{F}_{n-1}$ et $y \in \mathbb{F}_1 \cup \cdots \cup \mathbb{F}_{n-1} - \mathbb{F}_n$ et des combinaisons de la forme $x + \alpha y$). Quel énoncé peut-on obtenir lorsque $\mathbb{K}$ est un corps fini ?

DÉFINITION 1-2.10 (Famille génératrice) *Une famille $\mathcal{F}$ de vecteurs d'un e.v. $\mathbb{E}$ est dite génératrice de $\mathbb{E}$ si et seulement si*

$$\mathbb{E} = vect(\mathcal{F})$$

En particulier, une famille $\mathcal{G}$ non vide d'un espace vectoriel $\mathbb{E}$ est génératrice d'un s.e.v. $\mathbb{F}$ si et seulement si **tout vecteur de $\mathcal{G}$ appartient à $\mathbb{F}$** et si tout vecteur de $\mathbb{F}$ est combinaison linéaire de $\mathcal{G}$.

1-2.3 Familles équivalentes

DÉFINITION 1-2.11 *Deux familles (non vides) $\mathcal{F} = (x_i)_{i\in I}$ et $\mathcal{G} = (y_j)_{j\in J}$ de vecteurs de $\mathbb{E}$ sont dites équivalentes si et seulement elles engendrent le même sous-espace vectoriel de $\mathbb{E}$, soit*

$$vect(\mathcal{F}) = vect(\mathcal{G})$$

Il est clair que cette relation est réflexive, symétrique et transitive. La caractérisation des vecteurs de vect($\mathcal{F}$) comme combinaisons linéaires des vecteurs de $\mathcal{F}$ et la proposition **1-1**.11 donnent immédiatement :

THÉORÈME 1-2.12 Deux familles (non vides) $\mathcal{F} = (x_i)_{i\in I}$ et $\mathcal{G} = (y_j)_{j\in J}$ de vecteurs de $\mathbb{E}$ sont équivalentes si et seulement tout vecteur de $\mathcal{F}$ est combinaison linéaire de $\mathcal{G}$ et réciproquement.

En particulier, on transforme une famille $\mathcal{F} = (x_i)_{i\in I}$ en une famille équivalente en opérant une des "manipulations élémentaires" :

- Ajouter à un des vecteurs une combinaison linéaire des autres, ce qu'on symbolise par

$$x_k \leftarrow x_k + \sum_{i\neq k} \alpha_i x_i$$

(où les α_i sont presque tous nuls si I est infini)

- Multiplier un des vecteurs par un scalaire non nul

$$x_k \leftarrow \alpha x_k$$

- Permuter deux vecteurs de la famille

$$x_k \leftrightarrow x_l$$

ce qui revient formellement à composer l'application $i \mapsto x_i$ avec la transposition d'indices (k, l).

1-2.4 Somme de sous-espaces vectoriels

DÉFINITION 1-2.13 *Si $\mathbb{F}_1, \ldots, \mathbb{F}_p$ sont p sous-espaces vectoriels d'un e.v. $\mathbb{E}$, on appelle somme des sous-espaces $(\mathbb{F}_i)_{1\leqslant i\leqslant p}$ le sous-espace engendré par $\bigcup_{i=1}^{p} \mathbb{F}_i$. C'est donc le plus petit sous-espace de $\mathbb{E}$ (pour l'inclusion) contenant tous les $\mathbb{F}_i$. On note*

$$\mathbb{F}_1 + \cdots + \mathbb{F}_p = \sum_{i=1}^{p} \mathbb{F}_i = vect\left(\bigcup_{i=1}^{p} \mathbb{F}_i\right)$$

Comme pour le théorème **1-2**.6 (la démonstration est analogue), on a la caractérisation des vecteurs de la somme $\mathbb{F}_1 + \cdots + \mathbb{F}_p$:

PROPOSITION 1-2.14 La somme $\mathbb{F}_1 + \cdots + \mathbb{F}_p$ est exactement l'ensemble des vecteurs x de $\mathbb{E}$ admettant (au moins) une écriture sous la forme

$$x = x_1 + \cdots + x_p \textbf{ avec } \forall i \quad x_i \in \mathbb{F}_i$$

EXERCICE 1-2.15 Montrer que si $\mathcal{A}_1, \ldots, \mathcal{A}_p$ sont des familles de vecteurs engendrant respectivement $\mathbb{F}_1, \ldots, \mathbb{F}_p$, $\mathcal{A}_1 \cup \ldots \cup \mathcal{A}_p$ engendre $\mathbb{F}_1 + \cdots + \mathbb{F}_p$.

REMARQUE 1-2.16 Si $(\mathbb{F}_i)_{i\in I}$ est une famille infinie de s.e.v. de $\mathbb{E}$, on peut encore définir

$$\sum_{i\in I} \mathbb{F}_i = \text{vect}\left(\bigcup_{i\in I} \mathbb{F}_i\right)$$

Il est alors facile de montrer que les éléments de $\sum_{i\in I} \mathbb{F}_i$ sont exactement les vecteurs admettant une décomposition sous la forme $\sum_{i\in I} x_i$ où les $x_i \in \mathbb{F}_i$ sont presque tous nuls.

On va maintenant s'intéresser à l'unicité de la décomposition obtenue par la proposition **1-2**.14, en étudiant d'abord le cas des droites vectorielles. C'est la notion de famille libre ou liée, et plus généralement de sous-espaces indépendants.

1-3 Dépendance et indépendance linéaire

1-3.1 Famille libre, famille liée

Soit $\mathcal{A} = (x_i)_{i\in I}$ une famille non vide de vecteurs d'un e.v. $\mathbb{E}$. Si $j \in I$, on note $\mathcal{A}_j = (x_i)_{i\in I-\{j\}}$ la famille obtenue en retirant le vecteur x_j de la famille $\mathcal{A}$ (attention : on peut cependant peut-être retrouver ce vecteur dans $\mathcal{A}_j$ si l'application $i \mapsto x_i$ n'est pas injective). On note $\mathbb{F}$ =vect($\mathcal{A}$) et $\mathbb{F}_j$=vect($\mathcal{A}_j$). On a évidemment

$$\mathbb{F}_j \subset \mathbb{F}$$

Deux possibilités s'excluent mutuellement :

1. Pour tout $j \in I$, $\mathbb{F}_j$ est strictement inclus dans $\mathbb{F}$. Il est alors clair que toute sous-famille (stricte) de $\mathcal{A}$ engendre un sous-espace strictement inclus dans $\mathbb{F}$ (puisqu'une telle sous-famille est incluse dans une des $\mathcal{A}_j$) : $\mathcal{A}$ est **famille génératrice minimale** de $\mathbb{F}$ =vect($\mathcal{A}$). On dit alors que $\mathcal{A}$ est une **famille libre**, ou que **les vecteurs de $\mathcal{A}$ sont (linéairement) indépendants.**

2. Il existe $j \in I$ avec $\mathbb{F}_j = \mathbb{F}$. Si I est de cardinal au moins égal à 2, ceci équivaut à dire que x_j est combinaison linéaire des vecteurs de la famille $(x_i)_{i \in I - \{j\}} = \mathcal{A}_j$. En effet, si $\mathbb{F}_j = \mathbb{F}$, le vecteur x_j, qui est dans $\mathbb{F}$, est dans $\mathbb{F}_j$ et est donc combinaison linéaire de $\mathcal{A}_j$. Réciproquement, si le vecteur $x_j \in$vect$(\mathcal{A}_j) = \mathbb{F}_j$, $\mathbb{F}_j$ est un sous-espace de $\mathbb{E}$ qui contient tous les vecteurs de $\mathcal{A}$, donc qui contient vect$(\mathcal{A}) = \mathbb{F}$.

   $\mathcal{A}$ non génératrice minimale de vect($\mathcal{A}$) $\iff$
   il existe (au moins) un vecteur de $\mathcal{A}$ combinaison linéaire des autres.

   On dit alors que $\mathcal{A}$ est **une famille liée, les vecteurs de $\mathcal{A}$ sont dépendants.** (Si $\mathcal{A} = (x)$ ne contient qu'un vecteur, vect$(\mathcal{A})$ =vect$(\emptyset) = \{0_{\mathbb{E}}\}$ si et seulement si $x = 0_{\mathbb{E}}$).

Résumons la discussion précédente par la définition :

DÉFINITION 1-3.1 *Une famille $\mathcal{A} = (x_i)_{i \in I}$ de vecteurs est liée si et seulement si elle est réduite au vecteur nul ou si l'un des vecteurs est combinaison linéaire des autres. Dans ce cas, il existe une sous-famille stricte de $\mathcal{A}$ qui engendre vect($\mathcal{A}$). $\mathcal{A}$ est libre dans le cas contraire, donc si elle est génératrice minimale de vect($\mathcal{A}$).*

REMARQUE 1-3.2 Il est clair qu'une famille non injective est liée.

Attention ! Si $\mathcal{A}$ est liée, on ne peut choisir "au hasard" un vecteur qui serait combinaison linéaire des autres. Par exemple, si x et y sont deux vecteurs indépendants, la famille $(x_i)_{1 \leqslant i \leqslant 3}$ avec $x_1 = x$, $x_2 = -x$ et $x_3 = y$ est liée, mais x_3 n'est pas combinaison linéaire de x_1 et x_2.

1-3.2 Relations de liaison

On obtient la caractérisation commode de la dépendance linéaire :

PROPOSITION 1-3.3 La famille $\mathcal{A} = (x_i)_{i \in I}$ est liée s'il existe une famille de scalaires $(\alpha_i)_{i \in I}$ non tous nuls (mais bien entendu presque tous nuls si I est infini) tels que

$$\sum_{i \in I} \alpha_i x_i = 0_{\mathbb{E}}$$

Une telle égalité sera dite relation de liaison entre les vecteurs $(x_i)_{i \in I}$.

Démonstration : Si une telle relation existe, on choisit un indice i_0 avec $\alpha_{i_0} \neq 0$ et on obtient

$$x_{i_0} = \sum_{i \neq i_0} -\frac{\alpha_i}{\alpha_{i_0}} x_i$$

ce qui montre que $\mathcal{A}$ est liée (on notera que le fait que $\mathbb{K}$ soit un corps intervient dans cette démonstration : α_{i_0} est non nul, donc inversible pour la multiplication). Réciproquement, s'il existe un vecteur x_{i_0} combinaison linéaire des autres, soit

$$x_{i_0} = \sum_{i \neq i_0} \beta_i x_i$$

on a immédiatement une relation de liaison (à coefficients non tous nuls)

$$\sum_{i \in I} \alpha_i x_i = 0_{\mathbf{E}}$$

avec $\alpha_i = \beta_i$ pour $i \neq i_0$ et $\alpha_{i_0} = -1$. ∎

Comme une relation de liaison ne fait intervenir qu'un nombre fini de vecteurs, on en déduit immédiatement le corollaire :

COROLLAIRE 1-3.4 Une famille infinie de vecteurs est liée ssi il en existe une sous-famille finie qui est liée. Une famille est libre ssi toutes ses sous familles finies le sont.

COROLLAIRE 1-3.5 Toute sous-famille d'une famille libre est libre. Toute famille contenant une famille liée est liée.

EXERCICE 1-3.6 Si $(P_i)_{i \in I}$ est une famille de polynômes non nuls de $\mathbb{K}[X]$ avec pour $i \neq j$, $d^\circ P_i \neq d^\circ P_j$, cette famille est libre.

1-3.3 Indépendance linéaire dans des espaces fonctionnels

On est souvent amené à étudier l'indépendance linéaire de familles de fonctions dans des espaces tels que $\mathbb{R}^{\mathbb{R}}$, $\mathbb{R}^{[a,b]}$, $\mathbb{C}^{[a,b]}$ etc... Il faut alors penser à utiliser des techniques d'analyse (étude du comportement au voisinage d'un point, développements limités, comportement à l'infini, dérivation, ... si les fonctions considérées le permettent).

EXEMPLE 1-3.7 Etudier la dépendance linéaire dans $\mathbb{R}^{\mathbb{R}}$ de la famille de fonctions $(f_{n,\alpha})_{n \in \mathbb{N},\, \alpha \in \mathbb{R}}$ définies par

$$f_{n,\alpha}(x) = x^n e^{\alpha x}$$

(Indication : on considère une sous-famille finie $(f_{n_i,\alpha_i})_{1\leqslant i\leqslant p}$ extraite de cette famille, et on suppose l'existence d'une relation de liaison entre ces fonctions

$$\sum_{i=1}^{p} \lambda_i f_{n_i,\alpha_i} = 0$$

soit

$$\forall x \in \mathbb{R} \quad \sum_{i=1}^{p} \lambda_i x^{n_i} e^{\alpha_i x} = 0$$

Ce qui différencie nettement les fonctions f_{n_i,α_i}, c'est le comportement à l'infini. On peut (par exemple) multiplier par la fonction $x \mapsto e^{-\alpha x}$ où α est le plus grand des α_i et regarder le comportement pour $x \to +\infty$).

EXERCICE 1-3.8 Etudier de même la dépendance des $(f_{n,\alpha})_{n\in\mathbb{N},\alpha\in\mathbb{R}}$ dans $\mathbb{R}^{[0,1]}$, des $(f_{n,\alpha})_{n\in\mathbb{N},\alpha\in\mathbb{C}}$ dans $\mathbb{C}^{\mathbb{R}}$.

EXERCICE 1-3.9 Etudier la famille $(f_a)_{a\in\mathbb{R}}$ avec $f_a(x) = |x-a|$ dans $\mathbb{R}^{\mathbb{R}}$, dans $\mathbb{R}^{[-1,1]}$. (Ici, on peut utiliser la non dérivabilité de f_a en a).

1-3.4 Base d'un espace. Coordonnées d'un vecteur

DÉFINITION 1-3.10 *Une famille non vide $\mathcal{B} = (u_i)_{i\in I}$ de vecteurs d'un espace vectoriel $\mathbb{E}$ est dite base de $\mathbb{E}$ ssi $\mathcal{B}$ est libre et engendre $\mathbb{E}$.*

THÉORÈME 1-3.11 Une famille non vide $\mathcal{B} = (u_i)_{i\in I}$ de vecteurs de $\mathbb{E}$ est une base de $\mathbb{E}$ si et seulement si $\mathcal{B}$ est une famille libre maximale de $\mathbb{E}$ (ou, ce qui revient au même, est une famille génératrice minimale de $\mathbb{E}$).

Démonstration : Si $\mathcal{B}$ est une base, elle est libre. Tout vecteur x de $\mathbb{E}$ étant combinaison linéaire des vecteurs de $\mathcal{B}$ (génératrice), la famille obtenue en ajoutant x à $\mathcal{B}$ est liée, donc $\mathcal{B}$ est maximale parmi les familles libres de $\mathbb{E}$. Réciproquement, si $\mathcal{B}$ est libre maximale, la famille obtenue en ajoutant x quelconque de $\mathbb{E}$ à $\mathcal{B}$ est liée, ce qui assure l'existence d'une relation de liaison (à coefficients non tous nuls) de la forme

$$\lambda x + \sum_{i\in I} \alpha_i u_i = 0_{\mathbb{E}}$$

On a forcément $\lambda \neq 0$ (car on aurait sinon une relation de liaison non triviale entre les vecteurs de la famille libre $\mathcal{B}$), et on obtient x comme combinaison linéaire de $\mathcal{B}$

$$x = -\sum_{i\in I} \frac{\alpha_i}{\lambda} u_i$$

ce qui montre que $\mathcal{B}$ est aussi génératrice de $\mathbb{E}$. ■

COROLLAIRE 1-3.12 Une famille $\mathcal{B} = (u_i)_{i\in I}$ de vecteurs de $\mathbb{E}$ est une base de $\mathbb{E}$ si et seulement si tout vecteur x de $\mathbb{E}$ s'écrit de manière unique comme combinaison linéaire des vecteurs de $\mathcal{B}$

$$\forall x \in \mathbb{E} \quad \exists! \ (\alpha_i)_{i\in I} \in \mathbb{K}^I \ \textbf{(presque tous nuls)} \quad x = \sum_{i\in I} \alpha_i u_i$$

La famille $(\alpha_i)_{i\in I}$ est appelée famille des coordonnées du vecteur x dans la base $\mathcal{B}$.

Démonstration : Si $\mathcal{B}$ possède cette propriété, elle est évidemment génératrice. Une relation de liaison entre les vecteurs de $\mathcal{B}$ est en fait une écriture de $0_{\mathbb{E}}$ comme combinaison linéaire de $\mathcal{B}$. Comme cette écriture est unique, c'est forcément la relation triviale où tous les coefficients sont nuls : $\mathcal{B}$ est aussi libre. Réciproquement, si $\mathcal{B}$ est une base, tout vecteur de $\mathbb{E}$ est combinaison linéaire de $\mathcal{B}$ (génératrice). Si un vecteur x admet deux écritures

$$x = \sum_{i\in I} \alpha_i u_i = \sum_{i\in I} \beta_i u_i$$

on obtient la relation de liaison

$$\sum_{i\in I} (\alpha_i - \beta_i)\, u_i = 0_{\mathbb{E}}$$

et comme $\mathcal{B}$ est libre, on a

$$\forall i \in I \quad \alpha_i = \beta_i$$

La décomposition de x est donc unique. ■

COROLLAIRE 1-3.13 Si un espace vectoriel $\mathbb{E}$ possède une base $\mathcal{B} = (u_i)_{i\in I}$, on note $\mathbb{K}^{(I)}$ l'ensemble des familles indexées par I de scalaires presque tous nuls. L'application

$$\mathbb{K}^{(I)} \to \mathbb{E} \qquad (\alpha_i)_{i\in I} \mapsto \sum_{i\in I} \alpha_i u_i$$

est une bijection. $\mathbb{K}^{(I)}$ peut évidemment être muni d'une structure d'espace vectoriel (sous-espace vectoriel de $\mathbb{K}^I$) et l'application précédente est linéaire (voir plus loin). C'est un isomorphisme d'espaces vectoriels.

EXEMPLE 1-3.14 La famille $(X^n)_{n\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$, appelée base canonique de $\mathbb{K}[X]$.

1-3.5 Exemple : Fonctions polynômes

DÉFINITION 1-3.15 *Si $\mathbb{K}$ est un corps commutatif et p est un entier naturel, une application*

$$m : \mathbb{K}^p \to \mathbb{K}$$

est dite monomiale (à p variables) s'il existe un multi-indice

$$\alpha = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^p$$

tel que

$$\forall\, x = (x_1, \dots, x_p) \in \mathbb{K}^p \quad m(x) = x_1^{\alpha_1} \cdots x_p^{\alpha_p} \underset{déf.}{=} x^\alpha$$

Une fonction $P : \mathbb{K}^p \to \mathbb{K}$ est dite polynomiale (à p variables) ssi elle est combinaison linéaire de fonctions monômes (à p variables). L'ensemble des fonctions polynômes à p variables est donc le sous-espace vectoriel de $\mathbb{K}^{\mathbb{K}^p}$ engendré par les fonctions monômes. Si $\alpha = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^p$ est un multi-indice, on note $|\alpha| = \alpha_1 + \cdots + \alpha_p$ ("longueur de α"). Comme une fonction polynôme ne fait intervenir qu'un nombre fini de monômes, pour toute telle fonction P il existera un entier m (dépendant de P) tel que

$$\forall\, x \in \mathbb{K}^p \quad P(x) = \sum_{\substack{\alpha \in \mathbb{N}^p \\ |\alpha| \leqslant m}} \lambda_\alpha x^\alpha$$

où les λ_α sont des scalaires.

THÉORÈME 1-3.16 Si $\mathbb{K}$ est un corps infini (ceci est donc valable dans les cas usuels $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), les fonctions monômes forment une base de l'espace des fonctions polynômes à p variables sur $\mathbb{K}$, base qu'on appelle base canonique de cet espace vectoriel.

Démonstration : La démonstration se fait par récurrence sur p. Remarquons d'abord que, par définition, les fonctions monômes engendrent l'espace considéré. Pour $p = 1$, l'indépendance linéaire des fonctions monômes traduit exactement le fait qu'un polynôme fonctionnellement nul est formellement nul (car un polynôme de degré m s'annule au maximum en m points distincts de $\mathbb{K}$). Si on suppose qu'au rang p les fonctions monômes sont indépendantes, considérons une relation de liaison entre fonctions monômes à $p+1$ variables. On obtient un entier m et des scalaires λ_α tels que

$$\forall\, x \in \mathbb{K}^{p+1} \quad \sum_{\substack{\alpha \in \mathbb{N}^{p+1} \\ |\alpha| \leqslant m}} \lambda_\alpha x^\alpha = 0$$

En identifiant le $p+1$-uplet $(\alpha_1, \dots, \alpha_p, \alpha_{p+1})$ avec le couple (β, α_{p+1}) où $\beta = (\alpha_1, \dots, \alpha_p) \in \mathbb{K}^p$, ceci peut s'écrire également

$$\forall\, y \in \mathbb{K}^p \quad \forall\, z \in \mathbb{K} \quad \sum_{i=0}^{m} \left(\sum_{\substack{\beta \in \mathbb{N}^p \\ |\beta| \leqslant m-i}} \lambda_{(\beta,i)} y^\beta \right) z^i = 0$$

En prenant y arbitraire et en utilisant le résultat pour $p = 1$, on obtient

$$\forall y \in \mathbb{K}^p \quad \forall i \in [0, m] \quad \sum_{\substack{\beta \in \mathbb{N}^p \\ |\beta| \leqslant m-i}} \lambda_{(\beta,i)} y^\beta = 0$$

Il suffit alors d'appliquer l'hypothèse de récurrence pour montrer que tous les coefficients λ_α sont nuls. ■

REMARQUE 1-3.17 Lorsque $\mathbb{K}$ est un corps fini de cardinal k, le résultat précédent n'est plus valable. Déjà avec $p = 1$, si $\mathbb{K} = \{x_1, \ldots, x_k\}$, le polynôme

$$P(X) = \prod_{i=1}^{k} (X - x_i)$$

n'est pas formellement nul (il est de degré k et de coefficient dominant égal à 1), mais la fonction polynôme associée est identiquement nulle. On peut voir (*cf.* chapitre **17**) qu'on a en fait

$$P(X) = X^k - X$$

EXERCICE 1-3.18 Montrer que, si $\mathbb{K}$ est un corps fini, toute fonction $\mathbb{K}^p \to \mathbb{K}$ est polynomiale (on considèrera d'abord le cas $p = 1$, et on pourra utiliser, mais ce n'est nullement indispensable, les polynômes interpolateurs de Lagrange : voir chapitre sur la dualité).

1-3.6 Sous-espaces indépendants

On se limite ici aux familles finies de sous-espaces vectoriels d'un espace $\mathbb{E}$.

1-3.6.1 Définition

DÉFINITION 1-3.19 *Si $\mathbb{F}_1, \ldots, \mathbb{F}_p$ sont p sous-espaces vectoriels d'un espace $\mathbb{E}$, ils sont dits (linéairement) indépendants ssi l'application*

$$\prod_{i=1}^{p} \mathbb{F}_i \xrightarrow{\varphi} \sum_{i=1}^{p} \mathbb{F}_i \qquad (x_1, \ldots, x_p) \mapsto x_1 + \cdots + x_p$$

est bijective (elle est évidemment toujours surjective). Ceci équivaut à dire que tout vecteur x de $\sum_{i=1}^{p} \mathbb{F}_i$ admet une décomposition **unique** *sous la forme*

$$x = x_1 + \cdots + x_p \text{ avec } \forall i \quad x_i \in \mathbb{F}_i$$

On dit alors aussi que la somme des sous-espaces $\mathbb{F}_i$ est une **somme directe**, *et on la note*

$$\sum_{i=1}^{p} \mathbb{F}_i = \bigoplus_{i=1}^{p} \mathbb{F}_i$$

Si $x \in \bigoplus_{i=1}^{p} \mathbb{F}_i$ *se décompose comme précédemment,* x_j *est appelé* ***composante de*** $x \in \bigoplus_{i=1}^{p} \mathbb{F}_i$ ***selon le sous-espace*** $\mathbb{F}_j$.

REMARQUE 1-3.20 L'application φ considérée plus haut est toujours linéaire. Si les $\mathbb{F}_i$ sont indépendants, c'est un isomorphisme d'espaces vectoriels.

REMARQUE 1-3.21 Il est facile de voir que p vecteurs $e_1, \ldots, e_p$ sont indépendants ssi les droites vectorielles $\mathbb{D}_i$ =vect(e_i) sont indépendantes, puisque l'indépendance de ces vecteurs signifie que $(e_i)_{1 \leqslant i \leqslant p}$ est une base de

$$\text{vect}\,(e_1, \ldots, e_p) = \sum_{i=1}^{p} \mathbb{D}_i$$

donc que tout vecteur de la somme se décompose de manière unique dans la somme des $\mathbb{D}_i$.

1-3.6.2 Caractérisation

THÉORÈME 1-3.22 Deux sous-espaces $\mathbb{F}_1$ et $\mathbb{F}_2$ sont indépendants ssi

$$\mathbb{F}_1 \cap \mathbb{F}_2 = \{0_\mathbb{E}\}$$

Plus généralement p sous-espaces $\mathbb{F}_1, \ldots, \mathbb{F}_p$ sont indépendants ssi

$$\forall j \in \{1, \ldots, p\} \qquad \mathbb{F}_j \cap \left(\sum_{i \neq j} \mathbb{F}_i \right) = \{0_\mathbb{E}\}$$

Démonstration : Supposons les p sous-espaces $\mathbb{F}_1, \ldots, \mathbb{F}_p$ indépendants et démontrons la propriété. Comme l'addition est commutative dans $\mathbb{E}$, il n'est pas restrictif de n'envisager que le cas $j = p$, pour des raisons de commodité d'écriture. Soit $x \in \mathbb{F}_p \cap \left(\sum_{i=1}^{p-1} \mathbb{F}_i \right)$. Il admet alors une écriture sous la forme

$$x = x_1 + \cdots + x_{p-1} \text{ avec } x_i \in \mathbb{F}_i$$

et peut aussi s'écrire

$$x = x_1 + \cdots + x_{p-1} + 0_\mathbb{E} = 0_\mathbb{E} + \cdots + 0_\mathbb{E} + x$$

Par indépendance des sous-espaces $\mathbb{F}_i$, cette écriture est unique, ce qui donne bien $x = 0_\mathbb{E}$. Réciproquement, supposons la propriété

$$\forall j \in \{1, \ldots, p\} \qquad \mathbb{F}_j \cap \left(\sum_{i \neq j} \mathbb{F}_i \right) = \{0_E\}$$

et supposons qu'un vecteur x de la somme des $\mathbb{F}_i$ admette deux décompositions de la forme

$$x = \sum_{i=1}^{p} x_i = \sum_{i=1}^{p} y_i \text{ avec pour tout } i,\ x_i \text{ et } y_i \in \mathbb{F}_i$$

On a alors

$$\forall j \in \{1, \ldots, p\} \quad y_j - x_j = \sum_{i \neq j} (x_i - y_i) \in \mathbb{F}_j \cap \left(\sum_{i \neq j} \mathbb{F}_i \right) = \{0_E\}$$

ce qui donne $x_j = y_j$ et prouve l'indépendance des sous-espaces. ■

ATTENTION ! Deux erreurs fréquentes à ne pas commettre :

- Si $p \geqslant 3$, la condition démontrée plus haut entraîne bien

$$\text{Pour } i \neq j \quad \mathbb{F}_j \cap \mathbb{F}_i = \{0_E\}$$

(puisque $\mathbb{F}_i \subset \sum_{k \neq j} \mathbb{F}_k$) : l'indépendance d'une famille de sous-espaces entraîne l'indépendance "2 à 2" et plus généralement l'indépendance de toute famille extraite de la famille : c'est une conséquence immédiate de la définition, la restriction d'une injection à une partie de l'ensemble de départ est encore une injection. Mais la réciproque est évidemment fausse ! C'est la même chose que pour les systèmes de vecteurs où l'indépendance 2 à 2 n'entraîne pas la liberté du système total . Si e_1 et e_2 sont deux vecteurs libres, les sous-espaces

$$\mathbb{F}_1 = \text{vect}(e_1),\ \mathbb{F}_2 = \text{vect}(e_2) \text{ et } \mathbb{F}_3 = \text{vect}(e_1 + e_2)$$

sont deux à deux indépendants, mais ne sont pas indépendants.

- La caractérisation précédente n'est pas très commode d'utilisation, car elle oblige à faire la vérification **pour tous les indices** j. Une vérification avec un indice particulier est insuffisante, comme le montre l'exemple

$$\mathbb{F}_1 = \text{vect}(e_1),\ \mathbb{F}_2 = \text{vect}(e_1) \text{ et } \mathbb{F}_3 = \text{vect}(e_1 + e_2)$$

(avec toujours e_1 et e_2 indépendants). Ces sev ne sont évidemment pas indépendants bien que

$$\mathbb{F}_3 \cap (\mathbb{F}_1 + \mathbb{F}_2) = \{0_E\}$$

On peut cependant parfois utiliser le résultat suivant, où l'on prouve une espèce d'associativité :

PROPOSITION 1-3.23 Si $p \geqslant 3$ et $\mathbb{F}_1, \ldots, \mathbb{F}_p$ sont p sous-espaces vectoriels, ils sont indépendants ssi $\mathbb{F}_1, \ldots, \mathbb{F}_{p-1}$ le sont et

$$\mathbb{F}_p \cap \bigoplus_{i=1}^{p-1} \mathbb{F}_i = \{0_{\mathbb{E}}\}$$

Démonstration : Ces conditions sont évidemment nécessaires d'après ce qui précède. Supposons à présent qu'elles soient vérifiées. Si $x \in \sum_{i=1}^{p} \mathbb{F}_i$ admet deux écritures

$$x = \sum_{i=1}^{p} x_i = \sum_{i=1}^{p} y_i \text{ avec pour tout } i,\ x_i \text{ et } y_i \in \mathbb{F}_i$$

on a comme précédemment

$$y_p - x_p = \sum_{i=1}^{p-1} (x_i - y_i) \in \mathbb{F}_p \cap \left(\bigoplus_{i=1}^{p-1} \mathbb{F}_i \right) = \{0_E\}$$

donc $x_p = y_p$, puis, puisque $\mathbb{F}_1, \ldots, \mathbb{F}_{p-1}$ sont indépendants

$$\sum_{i=1}^{p-1} (x_i - y_i) = 0_{\mathbf{E}} \Rightarrow \forall i \quad x_i = y_i$$

ce qui prouve bien l'indépendance de $\mathbb{F}_1, \ldots, \mathbb{F}_p$. ■

1-3.6.3 Cas de sous-espaces dont on connaît des bases

Commençons par définir la **concaténation** de deux familles : il s'agit tout simplement de "juxtaposer" deux familles $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$. On pourrait penser indexer cette juxtaposition par $I \cup J$, mais il y a une difficulté si les ensembles I et J ne sont pas disjoints. C'est pour cela qu'on est amené à la définition plus formelle :

DÉFINITION 1-3.24 (Concaténation de familles) *Soient deux familles de* $\mathbb{E}$ $\mathcal{F} = (x_i)_{i \in I}$ *et* $\mathcal{G} = (y_j)_{j \in J}$. *On appelle concaténée de ces deux familles et on note* $\mathcal{F} \dot{+} \mathcal{G}$ *la famille indexée par* $K = (I \times \{1\}) \cup (J \times \{2\})$

$$\mathcal{F} \dot{+} \mathcal{G} = (z_k)_{k \in K}$$

avec, pour $i \in I$, $z_{(i,1)} = x_i$ *et, pour* $j \in J$, $z_{(j,2)} = y_j$.

On définirait de même la concaténation de p familles indexées par les ensembles $I_1, \ldots, I_p$, en travaillant avec les ensembles d'indices

$$I_1 \times \{1\}, \ldots, I_p \times \{p\}$$

THÉORÈME 1-3.25 Si $\mathbb{F}_1, \ldots, \mathbb{F}_p$ est une famille de p sous-espaces vectoriels possédant des bases respectives $\mathcal{B}_1, \ldots, \mathcal{B}_p$, ces sous-espaces sont indépendants ssi la famille obtenue par concaténation de ces bases est libre. Si cette propriété est vérifiée, cette réunion est alors une base de la somme (directe) des $\mathbb{F}_i$.

Démonstration : On sait déjà que, dans tous les cas, la famille $\mathcal{B}_1 \dot{+} \cdots \dot{+} \mathcal{B}_p$ engendre $\sum \mathbb{F}_i$. Un vecteur de la somme des $\mathbb{F}_i$ s'écrit

$$x = \sum_{i=1}^{p} x_i$$

avec $x_i \in \mathbb{F}_i$, x_i s'écrivant de manière unique comme combinaison linéaire des vecteurs de $\mathcal{B}_i$. Pour que la décomposition de x soit unique dans $\sum \mathbb{F}_i$, il est clair qu'il est nécessaire et suffisant que x admette une décomposition unique comme combinaison linéaire de $\mathcal{B}_1 \dot{+} \cdots \dot{+} \mathcal{B}_p$, ce qui traduit exactement la liberté de cette famille. ■

REMARQUE 1-3.26 Il nous arrivera souvent d'écrire $\mathcal{B}_1 \cup \mathcal{B}_2$ pour $\mathcal{B}_1 \dot{+} \mathcal{B}_2$. Il faut cependant garder à l'esprit qu'il ne s'agit pas d'une union ensembliste, et que, si un vecteur apparaît à la fois dans $\mathcal{B}_1$ et $\mathcal{B}_2$, il apparaît deux fois dans cette "union".

1-3.7 Sous-espaces supplémentaires

DÉFINITION 1-3.27 *Deux sous-espaces $\mathbb{F}_1$ et $\mathbb{F}_2$ d'un espace vectoriel* $\mathbf{E}$ *sont dits supplémentaires (dans* $\mathbb{E}$*) si et seulement si*

$$\mathbb{F}_1 \oplus \mathbb{F}_2 = \mathbb{E}$$

Ceci équivaut évidemment à dire que $\mathbb{F}_1 + \mathbb{F}_2 = \mathbb{E}$ et $\mathbb{F}_1 \cap \mathbb{F}_2 = \{0_E\}$, ou encore que tout vecteur de $\mathbb{E}$ admet une décomposition unique sous la forme $x_1 + x_2$ avec $x_i \in \mathbb{F}_i$. Lorsqu'il risque d'y avoir une confusion, il faut préciser l'espace ambiant. Par exemple, si $\mathbb{F}_1$ et $\mathbb{F}_2$ sont deux sev indépendants de $\mathbb{E}$, $\mathbb{F}_2$ est un supplémentaire de $\mathbb{F}_1$ **dans** $\mathbb{F}_1 \oplus \mathbb{F}_2$.

Il est très important d'avoir à l'esprit une représentation géométrique de cette situation : penser, dans l'espace où nous évoluons, à un plan dont un point O a été choisi comme origine (le "vecteur nul") pour se représenter $\mathbb{F}_1$. Pour se représenter $\mathbb{F}_2$, on considèrera n'importe quelle droite passant par O et non contenue dans $\mathbb{F}_1$. Tout point M de l'espace admet une projection M_1 sur $\mathbb{F}_1$ parallèlement à $\mathbb{F}_2$ et une projection M_2 sur $\mathbb{F}_2$ parallèlement à $\mathbb{F}_1$, ce qui correspond à une décomposition (unique)

$$\overrightarrow{OM} = \overrightarrow{OM_1} + \overrightarrow{OM_2}$$

où $\overrightarrow{OM_i}$ est dans (la direction de) $\mathbb{F}_i$(figure **1.1**). Cette vision géométrique élémentaire permettra d'éviter des erreurs, notamment de parler **du** supplémentaire d'un sous-espace $\mathbb{F}$ de $\mathbb{E}$. Si un supplémentaire existe, il est fort rare qu'il soit unique. Cela évitera aussi de confondre **supplémentaire** et **complémentaire**! Le complémentaire d'un sev de $\mathbb{E}$ n'est pas un sev (il ne contient déjà pas le vecteur nul). Si

$$\mathbb{F}_1 \oplus \mathbb{F}_2 = \mathbb{E}$$

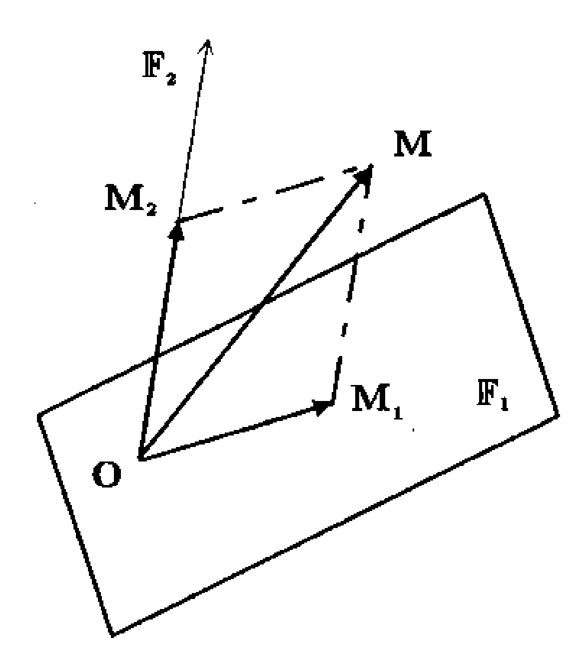


Figure **1.1** – *Sous-espaces supplémentaires*

cela ne signifie nullement $\mathbb{E} = \mathbb{F}_1 \cup \mathbb{F}_2$. L'espace où nous évoluons ne se réduit pas à un plan et une droite ! Ce qui est fondamental, c'est que l'on puisse géométriquement, et sans ambiguïté **reconstituer $\mathbb{E}$ à partir de $\mathbb{F}_1$** et $\mathbb{F}_2$.

Le résultat suivant n'est que la traduction dans un cas particulier du théorème 1-3.25 :

THÉORÈME 1-3.28 Deux sous-espaces $\mathbb{F}_1$ et $\mathbb{F}_2$ de $\mathbb{E}$ sont supplémentaires ssi la réunion d'une base de $\mathbb{F}_1$ et d'une base de $\mathbb{F}_2$ est une base de $\mathbb{E}$.

Il nous arrivera d'utiliser le théorème (admis), que nous démontrerons **dans le cadre des espaces de dimension finie** :

THÉORÈME 1-3.29 Dans un espace vectoriel $\mathbb{E}$, tout sous-espace $\mathbb{F}$ possède au moins un supplémentaire.

EXERCICE 1-3.30 En utilisant ce théorème, montrer que, si $\mathbb{F}$ est un sev de $\mathbb{E}$ ne contenant pas un vecteur x, $\mathbb{F}$ possède un supplémentaire contenant x. (Une petite réflexion **en dimension 3** peut amener à considérer le sev $\mathbb{G} = \mathbb{F} + \text{vect}(x)$).

EXERCICE 1-3.31 Si $\mathbb{F}$ et $\mathbb{G}$ sont deux sev de $\mathbb{E}$, si $\mathbb{F}_1$ est un supplémentaire de $\mathbb{F} \cap \mathbb{G}$ dans $\mathbb{F}$ et $\mathbb{G}_1$ est un supplémentaire de $\mathbb{F} \cap \mathbb{G}$ dans $\mathbb{G}$, montrer que

$$\mathbb{F} + \mathbb{G} = \mathbb{F}_1 \oplus (\mathbb{F} \cap \mathbb{G}) \oplus \mathbb{G}_1$$

(On obtient ainsi une décomposition géométrique de $\mathbb{F} + \mathbb{G}$ intéressante pour étudier des problèmes faisant intervenir l'intersection $\mathbb{F} \cap \mathbb{G}$).

EXERCICE 1-3.32 Soit P un polynôme de $\mathbb{K}[X]$ de degré $p > 0$. On note (P) l'ensemble des multiples de P dans $\mathbb{K}[X]$ (c'est l'idéal engendré par P). Montrer que (P)

est un sev de $\mathbb{K}[X]$ et que

$$\mathbb{K}[X] = (P) \oplus \mathbb{K}_{p-1}[X]$$

1-3.8 Projecteurs

1-3.8.1 Définition

DÉFINITION 1-3.33 *Si $\mathbb{E} = \mathbb{F}_1 \oplus \mathbb{F}_2$ est une décomposition d'un ev en deux sous-espaces supplémentaires, tout vecteur x de $\mathbb{E}$ possède une unique décomposition sous la forme*

$$x = x_1 + x_2 \text{ avec } x_i \in \mathbb{F}_i$$

On appelle projecteur sur $\mathbb{F}_1$ parallèlement à $\mathbb{F}_2$ l'application

$$p_1 : \mathbb{E} \to \mathbb{E} \quad x \mapsto p_1(x) = x_1$$

x_1 est la projection de x sur $\mathbb{F}_1$ parallèlement à $\mathbb{F}_2$. On définit de manière symétrique le projecteur p_2 sur $\mathbb{F}_2$ parallèlement à $\mathbb{F}_1$.

On verra ultérieurement qu'il s'agit d'endomorphismes de $\mathbb{E}$ qui vérifient évidemment

$$p_1 \circ p_1 = p_1, \quad p_2 \circ p_2 = p_2 \quad \text{et}$$
$$p_2 \circ p_1 = p_1 \circ p_2 = 0 \text{ (application identiquement nulle)}$$

1-3.8.2 Famille de projecteurs associés à une décomposition de l'espace

On généralise ici la définition précédente au cas où l'on a une décomposition de l'espace en somme de sous-espaces indépendants :

$$\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{F}_i$$

On sait alors (*cf.* proposition **1-3**.23) que pour $i \in \{1, \dots, p\}$

$$\mathbb{F}_i \text{ et } \bigoplus_{j \neq i} \mathbb{F}_j \text{ sont supplémentaires}$$

DÉFINITION 1-3.34 *On appelle famille de projecteurs associés à la décomposition $\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{F}_i$ la famille de projecteurs $(p_i)_{1 \leqslant i \leqslant p}$ où p_i est le projecteur sur $\mathbb{F}_i$ parallèlement à $\bigoplus_{j \neq i} \mathbb{F}_j$.*

p_i est donc l'application $\mathbb{E} \to \mathbb{E}$ qui à un vecteur x quelconque qui admet la décomposition

$$x = \sum_{j=1}^{p} x_j \text{ avec } x_j \in \mathbb{F}_j$$

associe sa composante $p_i(x) = x_i$ selon $\mathbb{F}_i$. On a encore évidemment

$$p_i \circ p_j = 0 \text{ si } i \neq j \quad \text{et} \quad p_i \circ p_i = p_i$$

1-4 Espaces vectoriels de dimension finie

On dit qu'un espace vectoriel $\mathbb{E}$ est **de dimension finie** s'il possède au moins une famille génératrice finie. Le résultat fondamental est alors l'existence de bases qui sont toutes finies et de même cardinal. Ce cardinal commun est appelé dimension de l'espace vectoriel $\mathbb{E}$.

1-4.1 Résultat préliminaire

Le théorème qui suit est à la base de la théorie de la dimension (intuitivement : le nombre de paramètres "indépendants" nécessaires pour repérer un vecteur).

THÉORÈME 1-4.1 Dans un espace vectoriel engendré par une famille de cardinal p, toute famille de cardinal strictement supérieur à p est liée.

Démonstration : Il suffit évidemment de démontrer que toute famille de $p+1$ vecteurs est liée. La démonstration se fait par récurrence sur p. Pour $\mathbb{E}_1 = \text{vect}(e_1)$ et $\mathcal{A}_1 = \{u_1, u_2\}$, on a l'existence de scalaires λ et μ tels que $u_1 = \lambda e_1$ et $u_2 = \mu e_1$. Si λ et μ sont tous deux nuls, $\mathcal{A}$ est liée. Sinon, on a la relation de liaison non triviale

$$\mu u_1 - \lambda u_2 = 0_{\mathbb{E}}$$

qui montre que, dans ce cas aussi, $\mathcal{A}$ est liée. Si on suppose le résultat démontré au rang p, on se place au rang $p+1$ en considérant $\mathbb{E}_{p+1} = \text{vect}(e_1, \ldots, e_{p+1})$ et $\mathcal{A}_{p+1} = \{u_1, \ldots, u_{p+2}\}$. Par hypothèse, il existe des scalaires $(\alpha_{i,j})_{\substack{1 \leqslant i \leqslant p+1 \\ 1 \leqslant j \leqslant p+2}}$ tels que

$$\forall j \in \{1, \ldots, p+2\} \quad u_j = \sum_{i=1}^{p+1} \alpha_{i,j} e_i$$

Si tous ces scalaires sont nuls, la famille $\mathcal{A}$ est évidemment liée. Sinon, en renumérotant les vecteurs des deux familles, on peut supposer $\alpha_{p+1,1} \neq 0$. On se ramène alors à considérer $p+1$ vecteurs dans $\mathbb{E}_p = \text{vect}(e_1, \ldots, e_p)$ en posant

$$\text{Pour } j \in \{2, \ldots, p+2\} \qquad v_j = \alpha_{p+1,1} u_j - \alpha_{p+1,j} u_1$$

(le vecteur e_{p+1} n'apparaît plus dans cette combinaison linéaire). Par hypothèse de récurrence, il existe une relation de liaison non triviale entre ces vecteurs de la forme

$$\sum_{j=2}^{p+2} \beta_j v_j = 0_{\mathbb{E}}$$

relation qui s'écrit aussi

$$\sum_{j=2}^{p+2} \beta_j \alpha_{p+1,1} u_j - \left(\sum_{j=2}^{p+2} \beta_j \alpha_{p+1} \right) u_1 = 0_{\mathbb{E}}$$

ce qui prouve que la famille $\mathcal{A}_{p+1}$ est liée puisqu'au moins un des coefficients $\beta_j \alpha_{p+1,1}$ est non nul. ∎

1-4.2 Existence de bases, dimension

THÉORÈME 1-4.2 (de la base incomplète) Si $\mathcal{G}$ est une famille génératrice d'un espace vectoriel $\mathbb{E}$ de dimension finie non réduit à $\{0_{\mathbb{E}}\}$, et si $\mathcal{L}$ est une famille libre de $\mathbb{E}$, il est possible de compléter $\mathcal{L}$ à l'aide de vecteurs (bien choisis !) de $\mathcal{G}$ pour obtenir une base de $\mathbb{E}$.

Démonstration : Remarquons d'abord que, dans ces hypothèses, $\mathcal{G}$ n'est pas nécessairement finie, mais le cardinal de $\mathcal{L}$ est inférieur ou égal à celui d'une famille génératrice finie de $\mathbb{E}$. La démonstration est basée sur le résultat suivant : si $\mathcal{F}$ est une famille libre et si x est un vecteur qui n'est pas combinaison linéaire de $\mathcal{F}$, la famille obtenue en rajoutant x à $\mathcal{F}$ est encore libre. Les espaces vect$(\mathcal{F})$ et vect(x) sont en effet indépendants et c'est le théorème **1-3**.25.

Si $\mathcal{L}$ engendre $\mathbf{E}$, c'est une base et il n'y a rien à faire. Sinon il existe au moins un vecteur x_1 de $\mathcal{G}$ qui n'est pas combinaison linéaire des vecteurs de $\mathcal{L}$: dans le cas contraire, on aurait

$$\mathbb{E} = \text{vect}\,(\mathcal{G}) \subset \text{vect}\,(\mathcal{L}) \subset \mathbb{E}$$

et $\mathcal{L}$ serait génératrice. On remplace alors la famille $\mathcal{L}$ par la famille (libre d'après la remarque précédente) $\mathcal{L}_1 = \mathcal{L} \cup (x)$ et on recommence le raisonnement... Au bout d'un nombre **fini** d'étapes, on arrive à une famille libre qui engendre $\mathbb{E}$ (puisque le cardinal d'une famille libre de $\mathbb{E}$ est majoré, le processus ne peut durer indéfiniment). ∎

COROLLAIRE 1-4.3 Soit E un espace vectoriel de dimension finie non réduit à $\{0_{\mathbb{E}}\}$. De toute famille génératrice $\mathcal{G}$ de $\mathbb{E}$ on peut extraire une base (qui contient un nombre fini d'éléments).

Démonstration : Il suffit d'appliquer le théorème précédent en partant de $\mathcal{L} = (x)$ où x est un vecteur non nul arbitraire extrait de $\mathcal{G}$. ∎

COROLLAIRE 1-4.4 (dimension) Un espace vectoriel $\mathbb{E}$ de dimension finie non réduit à $\{0_{\mathbb{E}}\}$ possède des bases qui sont toutes finies de même cardinal. Ce cardinal est appelé dimension de $\mathbb{E}$. Par convention, si $\mathbb{E}$ est réduit au vecteur nul, on dit que la dimension de $\mathbb{E}$ est nulle.

Démonstration : L'existence de bases, le fait que ces bases ont moins d'éléments qu'une famille génératrice finie de $\mathbb{E}$ ont déjà été vus. Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de cardinaux respectifs n_1 et n_2, comme $\mathcal{B}_1$ est génératrice et $\mathcal{B}_2$ libre, on a d'après le résultat préliminaire $n_2 \leqslant n_1$ et l'inégalité inverse est symétrique. ■

Dans un espace de dimension p, les familles libres ont au plus p éléments (résultat préliminaire), les familles génératrices ont au moins p éléments (s'il existait une famille génératrice de moins de p vecteurs, il ne pourrait y avoir de famille libre à p vecteurs).

EXERCICE 1-4.5 Montrer qu'un espace vectoriel $\mathbb{E}$ est de dimension finie ssi le cardinal d'une famille libre arbitraire de $\mathbb{E}$ est majoré.

Le théorème suivant est très souvent utilisé pour prouver qu'une famille est une base d'un espace vectoriel de dimension finie (dont on connaît la dimension) :

THÉORÈME 1-4.6 Si $\mathbb{E}$ est un espace vectoriel de dimension finie $p > 0$, une famille $\mathcal{B}$ de $\mathbb{E}$ est une base ssi elle est libre de cardinal p, ce qui équivaut aussi à dire que $\mathcal{B}$ est génératrice de cardinal p.

Démonstration : Ces conditions sont évidemment nécessaires pour que $\mathcal{B}$ soit une base (c'est la définition de la dimension). Si $\mathcal{B}$ est libre de cardinal p, toute sur-famille stricte de $\mathcal{B}$ est liée (car elle contient plus de p vecteurs). $\mathcal{B}$ est libre maximale, c'est donc une base. De même si $\mathcal{B}$ est génératrice de cardinal p, elle est génératrice minimale et est donc une base. ■

Rappelons pour terminer ce paragraphe que, si $\mathbb{E}$ possède une base

$$\mathcal{B} = (e_i)_{1 \leqslant i \leqslant p}$$

le choix de cette base induit une bijection

$$\varphi_{\mathcal{B}} : \mathbb{E} \to \mathbb{K}^p \qquad x = \sum_{i=1}^{p} \alpha_i e_i \mapsto \varphi_{\mathcal{B}}(x) = (\alpha_1, \ldots, \alpha_p)$$

qui est en fait un **isomorphisme d'espaces vectoriels**. C'est un cas particulier du corollaire **1-3**.13.

1-4.3 Dimension d'un s.e.v, rang d'un système de vecteurs

THÉORÈME 1-4.7 Si $\mathbb{F}$ est un sous-espace vectoriel d'un espace de dimension finie $\mathbb{E}$, $\mathbb{F}$ est de dimension finie avec $\dim \mathbb{F} \leqslant \dim \mathbb{E}$. **Il y a égalité ssi** $\mathbb{F} = \mathbb{E}$.

Démonstration : Si $\mathbb{F}$ est réduit à $\{0_{\mathbb{E}}\}$ il n'y a rien à faire. Sinon on considère dans $\mathbb{F}$ une famille libre de cardinal maximal. Une telle famille existe car une famille libre de $\mathbb{F}$ est aussi libre dans $\mathbb{E}$, donc est de cardinal inférieur ou égal à $\dim \mathbb{E}$. Cette famille est libre maximale dans $\mathbb{F}$. C'est donc une base de $\mathbb{F}$, qui est de dimension finie. On a bien $\dim \mathbb{F} \leqslant \dim \mathbb{E}$. S'il y a égalité, une base de $\mathbb{F}$ est libre dans $\mathbb{E}$ et de cardinal égal à $\dim \mathbb{E}$. C'est donc une base de $\mathbb{E}$ et $\mathbb{E} = \mathbb{F}$. ■

DÉFINITION 1-4.8 *Si $\mathcal{A}$ est une famille de vecteurs d'un espace vectoriel $\mathbb{E}$ (pas nécessairement de dimension finie), on dit que $\mathcal{A}$ est de rang fini ssi vect($\mathcal{A}$) est un sous-espace vectoriel de dimension finie. Dans ce cas, on définit le rang de $\mathcal{A}$ comme étant*

$$\operatorname{rang}(\mathcal{A}) = \dim \operatorname{vect}(\mathcal{A})$$

Si $\mathbb{E}$ est de dimension finie on a évidemment, pour toute famille de vecteurs de $\mathbb{E}$,

$$\operatorname{rang}(\mathcal{A}) \leqslant \dim \mathbb{E}$$

Il y a égalité ssi vect($\mathcal{A}$) $= \mathbb{E}$, c'est-à-dire si $\mathcal{A}$ engendre $\mathbb{E}$.

Si $\mathcal{A}$ est de cardinal fini, $\mathcal{A}$ étant génératrice de vect($\mathcal{A}$), on a évidemment

$$\operatorname{rang}(\mathcal{A}) \leqslant \operatorname{card}(\mathcal{A})$$

avec égalité ssi $\mathcal{A}$ est libre.

Si $\mathcal{A}$ est de rang fini, on peut appliquer à vect($\mathcal{A}$) le résultat du corollaire **1-4**.3. Si $\mathcal{A}$ n'est pas de rang fini, il n'existe pas de sous-famille finie extraite de $\mathcal{A}$ engendrant vect($\mathcal{A}$), et il est possible de construire de proche en proche, en utilisant la remarque à la base de la démonstration du théorème de la base incomplète **1-4**.2, une suite infinie de vecteurs indépendants extraite de $\mathcal{A}$. On a donc la caractérisation du rang :

THÉORÈME 1-4.9 Si $\mathcal{A}$ est de rang fini $r > 0$, r est le cardinal maximal d'une famille libre extraite de $\mathcal{A}$. Une telle famille de cardinal r est une base de vect($\mathcal{A}$), on dit que c'est une famille principale extraite de $\mathcal{A}$. Si $\mathcal{A}$ n'est pas de rang fini, il existe une suite de vecteurs libres extraite de $\mathcal{A}$.

Comme le rang d'un système de vecteurs ne dépend que du sous-espace engendré par ces vecteurs, deux familles équivalentes (*cf.* la section **1-2.3**) ont même rang. En particulier, on ne change pas le rang d'un système de vecteurs en transformant celui-ci par manipulations élémentaires.

La notion de famille triangulaire par rapport à une suite libre est derrière l'algorithme de recherche du rang par pivot de Gauss qu'on reverra dans le chapitre sur le calcul matriciel.

DÉFINITION 1-4.10 *Si $\mathcal{F} = (u_i)_{1\leqslant i\leqslant n}$ est une famille libre de vecteurs de $\mathbb{E}$, une famille de p vecteurs ($p \leqslant n$) $\mathcal{G} = (v_j)_{1\leqslant j\leqslant p}$ est dite triangulaire par rapport à la famille $\mathcal{F}$ ssi*

v_1 est non nul et colinéaire à u_1 et

$$\forall j \in \{2, \ldots, p\} \quad v_j \in \operatorname{vect}(u_1, \ldots, u_j) - \operatorname{vect}(u_1, \ldots, u_{j-1})$$

Il revient à la même chose de supposer l'existence de scalaires $(\alpha_{i,j})_{\substack{1\leqslant j\leqslant p\\1\leqslant i\leqslant j}}$ avec pour tout j $\alpha_{j,j} \neq 0$ tels que

$$\forall j \in \{1, \ldots, p\} \quad v_j = \sum_{i=1}^{j} \alpha_{i,j} u_i$$

THÉORÈME 1-4.11 Si $\mathcal{G} = (v_j)_{1\leqslant j\leqslant p}$ est triangulaire par rapport à une famille libre

$$\mathcal{F} = (u_i)_{1\leqslant i\leqslant n}$$

(avec $p \leqslant n$) alors $\mathcal{G}$ est libre (donc de rang p) et

$$\forall j \in \{1, \ldots, p\} \quad \operatorname{vect}(v_1, \ldots, v_j) = \operatorname{vect}(u_1, \ldots, u_j)$$

Démonstration : On a évidemment

$$\operatorname{vect}(v_1, \ldots, v_p) \subset \operatorname{vect}(u_1, \ldots, u_p)$$

Il suffit de démontrer que $\mathcal{G}$ est libre pour avoir

$$\dim \operatorname{vect}(v_1, \ldots, v_p) = p = \dim \operatorname{vect}(u_1, \ldots, u_p)$$

et avoir l'égalité de ces sev. En remplaçant ensuite dans le raisonnement p par j, le résultat en découle. Il est facile de voir que la résolution du système d'inconnues $\lambda_1, \ldots, \lambda_p$

$$\sum_{j=1}^{p} \lambda_j v_j = 0_{\mathbb{E}}$$

amène à l'écriture d'un système triangulaire dont la seule solution est la solution triviale (regarder d'abord la composante sur u_p pour obtenir $\lambda_p = 0$ etc...). ■

EXERCICE 1-4.12 Si $\mathcal{A}=(P_n)_{n\in\mathbb{N}}$ est une suite de polynômes de $\mathbb{K}[X]$, avec

$$\forall n \quad d^\circ P_n = n$$

montrer que $\mathcal{A}$ est une base de $\mathbb{K}[X]$

1-4.4 Application aux sous-espaces supplémentaires

THÉORÈME 1-4.13 Dans un espace vectoriel $\mathbb{E}$ de dimension finie, tout sous-espace vectoriel possède au moins un supplémentaire.

Démonstration : Notons $\mathbb{F}$ un sous-espace de $\mathbb{E}$. Si $\mathbb{F}$ est réduit à $\{0_{\mathbb{E}}\}$, $\mathbb{S} = \mathbb{E}$ est l'unique supplémentaire de $\mathbb{F}$. Le cas $\mathbb{F} = \mathbb{E}$ est symétrique. Sinon, si $\mathcal{B}_1$ est une base de $\mathbb{F}$, on peut (d'après le théorème de la base incomplète) compléter $\mathcal{B}_1$ (à l'aide de vecteurs d'une famille génératrice quelconque de $\mathbb{E}$) en une base de $\mathbb{E}$, soit $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$[3]. Le sous-espace $\mathbb{S}$ =vect($\mathcal{B}_2$) est un supplémentaire de $\mathbb{F}$ d'après le théorème **1-3**.25. ■

Lorsqu'on résout un problème d'algèbre linéaire, on est (parfois) amené à faire des calculs dans une base. Il est bien évident que le choix de cette base conditionne très souvent la lourdeur de ces calculs. C'est pourquoi il est indispensable de choisir judicieusement la base, pour qu'elle soit géométriquement adaptée au problème. Lorsque des sous-espaces interviennent de manière naturelle, on utilisera souvent les notions suivantes :

DÉFINITION 1-4.14 (bases adaptées) *Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}$, une base $\mathcal{B}$ de $\mathbb{E}$ est dite adaptée à $\mathbb{F}$ ssi $\mathcal{B}$ peut s'écrire $\mathcal{B}_1 \cup \mathcal{B}_2$ où $\mathcal{B}_1$ est une base de $\mathbb{F}$. De même, si*

$$\mathbb{E} = \mathbb{F}_1 \oplus \cdots \oplus \mathbb{F}_p$$

une base de $\mathbb{E}$ adaptée à cette décomposition de $\mathbb{E}$ en somme directe de sous-espaces indépendants est une base $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_p$ où chacune des $\mathcal{B}_i$ est une base de $\mathbb{F}_i$ (on suppose bien évidemment qu'aucun des sous-espaces n'est réduit au vecteur nul).

En dimension finie, le théorème de la base incomplète assure toujours l'existence de telles bases. En dimension infinie, ce résultat est encore vrai pour peu que l'on admette l'existence de bases pour tout espace vectoriel non réduit au vecteur nul et le théorème **1-3**.29.

1-4.5 Dimension d'une somme, d'un produit

THÉORÈME 1-4.15 Si $\mathbb{F}$ et $\mathbb{G}$ sont deux sous-espaces d'un espace $\mathbb{E}$, $\mathbb{F} + \mathbb{G}$ est de dimension finie si et seulement si $\mathbb{F}$ et $\mathbb{G}$ le sont et on a alors

$$\dim(\mathbb{F} + \mathbb{G}) = \dim \mathbb{F} + \dim \mathbb{G} - \dim(\mathbb{F} \cap \mathbb{G})$$

Démonstration : $\mathbb{F}$ et $\mathbb{G}$ sont deux sous-espaces vectoriels de $\mathbb{F} + \mathbb{G}$, et la réunion d'une base de $\mathbb{F}$ et d'une base de $\mathbb{G}$ engendrant $\mathbb{F} + \mathbb{G}$, il est clair que $\mathbb{F} + \mathbb{G}$ est de dimension finie si et seulement si $\mathbb{F}$ et $\mathbb{G}$ le sont. Si c'est le cas, la formule donnant la dimension de $\mathbb{F} + \mathbb{G}$ est évidente si $\mathbb{F}$ et $\mathbb{G}$ sont indépendants (théorème **1-3**.25). Si ce n'est

[3] Voir remarque **1-3**.26

pas le cas, on considère un supplémentaire $\mathbb{F}'$ de $\mathbb{F} \cap \mathbb{G}$ dans $\mathbb{F}$ (existe bien puisque $\mathbb{F}$ est de dimension finie). On a alors

$$\mathbb{F} + \mathbb{G} = (\mathbb{F}' + \mathbb{F} \cap \mathbb{G}) + \mathbb{G} = \mathbb{F}' + (\mathbb{F} \cap \mathbb{G} + \mathbb{G}) = \mathbb{F}' \oplus \mathbb{G}$$

la dernière somme étant directe, puisque

$$\mathbb{F}' \cap \mathbb{G} = (\mathbb{F}' \cap \mathbb{F}) \cap \mathbb{G} = \mathbb{F}' \cap (\mathbb{F} \cap \mathbb{G}) = \{0_{\mathbb{E}}\}$$

On a donc

$$\dim(\mathbb{F} + \mathbb{G}) = \dim \mathbb{F}' + \dim \mathbb{G}$$

et on en déduit le résultat puisque $\dim \mathbb{F} = \dim \mathbb{F}' + \dim(\mathbb{F} \cap \mathbb{G})$. ■

COROLLAIRE 1-4.16 Si $\mathbb{E}$ et $\mathbb{F}$ sont deux espaces vectoriels, $\mathbb{E} \times \mathbb{F}$ est de dimension finie ssi $\mathbb{E}$ et $\mathbb{F}$ le sont et on a alors

$$\dim(\mathbb{E} \times \mathbb{F}) = \dim \mathbb{E} + \dim \mathbb{F}$$

Démonstration : Il suffit de remarquer que

$$\mathbb{E} \times \mathbb{F} = (\mathbb{E} \times \{0_{\mathbb{F}}\}) \oplus (\{0_{\mathbb{E}}\} \times \mathbb{F})$$

Il est clair que, si $(u_i)_{i \in I}$ et $(v_j)_{j \in J}$ sont des bases respectives de $\mathbb{E}$ et $\mathbb{F}$, la famille

$$\mathcal{B} = ((u_i, 0_{\mathbb{F}}))_{i \in I} \cup ((0_{\mathbb{E}}, v_j))_{j \in J}$$

est une base de $\mathbb{E} \times \mathbb{F}$. ■

Du théorème précédent découle une caractérisation intéressante des sous-espaces supplémentaires **en dimension finie** :

COROLLAIRE 1-4.17 Si $\mathbb{E}$ est un espace vectoriel de dimension finie, $\mathbb{F}$ et $\mathbb{G}$ sous-espaces de $\mathbb{E}$ sont supplémentaires ssi

$$\mathbb{E} = \mathbb{F} + \mathbb{G} \quad \textbf{et} \quad \dim \mathbb{E} = \dim \mathbb{F} + \dim \mathbb{G}$$

ou encore si et seulement si

$$\mathbb{F} \cap \mathbb{G} = \{0_{\mathbb{E}}\} \quad \textbf{et} \quad \dim \mathbb{E} = \dim \mathbb{F} + \dim \mathbb{G}$$

On a également la généralisation suivante de ce corollaire :

THÉORÈME 1-4.18 Si $(\mathbb{F}_i)_{1 \leqslant i \leqslant p}$ sont p sous-espaces de dimension finie d'un espace vectoriel $\mathbb{E}$, ces sous-espaces sont indépendants si et seulement si

$$\dim \sum_{i=1}^{p} \mathbb{F}_i = \sum_{i=1}^{p} \dim \mathbb{F}_i$$

Démonstration : Remarquons d'abord que les hypothèses assurent que la somme des $\mathbb{F}_i$ est de dimension finie. La condition est évidemment nécessaire. Montrons qu'elle est suffisante. Si les sous-espaces n'étaient pas indépendants, il existerait un indice pour lequel

$$\mathbb{F}_j \cap \left(\sum_{i \neq j} \mathbb{F}_i\right) \neq \{0_E\}$$

On aurait alors

$$\dim \sum_{i=1}^{p} \mathbb{F}_i = \dim \left(\mathbb{F}_j + \left(\sum_{i \neq j} \mathbb{F}_i\right)\right) =$$
$$\dim \mathbb{F}_j + \dim \left(\sum_{i \neq j} \mathbb{F}_i\right) - \dim \left(\mathbb{F}_j \cap \left(\sum_{i \neq j} \mathbb{F}_i\right)\right)$$

ce qui donne

$$\dim \sum_{i=1}^{p} \mathbb{F}_i < \dim \mathbb{F}_j + \dim \left(\sum_{i \neq j} \mathbb{F}_i\right) \leqslant \dim \mathbb{F}_j + \sum_{i \neq j} \dim (\mathbb{F}_i)$$

puisque la dimension d'une somme est toujours inférieure ou égale à la somme des dimensions (la réunion des bases des sev est génératrice de la somme des sev). Ce qui nie la propriété supposée. ■

1-4.6 Restriction du corps des scalaires

EXERCICE 1-4.19 Soit $\mathbb{E}$ un $\mathbb{K}$-espace vectoriel de dimension finie non réduit à $\{0_{\mathbb{E}}\}$ et $\mathbb{K}_1$ un sous-corps de $\mathbb{K}$. Pour que, par restriction du corps des scalaires, $\mathbb{E}$ soit un $\mathbb{K}_1$-ev de dimension finie, il faut et il suffit que $\mathbb{K}$ soit un $\mathbb{K}_1$-ev de dimension finie et on a alors

$$\dim_{\mathbb{K}_1} \mathbb{E} = \dim_{\mathbb{K}} \mathbb{E} \times \dim_{\mathbb{K}_1} \mathbb{K}$$

Indication : Considérer d'abord le cas (le plus important pour nous) $\mathbb{K} = \mathbb{C}$ et $\mathbb{K}_1 = \mathbb{R}$. Montrer qu'alors, si $(u_1, \ldots, u_p)$ est une base du $\mathbb{C}$-ev $\mathbb{E}$, $(u_1, iu_1, \ldots, u_p, iu_p)$ est une base de $\mathbb{E}$ considéré comme $\mathbb{R}$-ev : la dimension d'un espace réel déduit d'un espace complexe est en particulier toujours paire. Dans le cas général, si $(u_k)_{k \in K}$ et $(\alpha_j)_{j \in J}$ sont des bases respectives de $\mathbb{E}$ sur $\mathbb{K}$ et de $\mathbb{K}$ sur $\mathbb{K}_1$, montrer que $(\alpha_j u_k)_{(k,j) \in K \times J}$ est une base de $\mathbb{E}$ sur $\mathbb{K}_1$.

1-5 Exercices

EXERCICE 1-5.1 Soit $(x_i)_{1\leq i\leq n}$ une famille libre d'un K-ev et $y = \sum_{i=1}^{n} \alpha_i x_i$. Condition nécessaire et suffisante sur les α_i pour que $(y + x_i)_{1\leq i\leq n}$ soit libre.

EXERCICE 1-5.2 Soient $x_1 < x_2 < \cdots < x_n$ des réels. On pose $x_0 = -\infty$ et $x_{n+1} = +\infty$. E est l'ensemble des fonctions $\mathbf{R} \to \mathbf{R}$, de classe C^1, dont la restriction à chaque $]x_i, x_{i+1}[$ est un polynôme de degré $\leqslant 2$. Montrer que E est un espace vectoriel. En donner la dimension et une base.

EXERCICE 1-5.3 Soit E un ev de dimension n. On considère une famille de vecteurs non nuls $\{x_1, \cdots, x_p\}$ de rang r. On en extrait $q \geqslant 1$ vecteurs et on obtient une famille de rang r'. Trouver un minorant de r'.

EXERCICE 1-5.4 Dans $\mathbf{R}^{[0,1]}$ le système $(f_a)_{a\in\mathbf{R}^{+*}}$ défini par $f_a(t) = \sqrt{a+t}$ est-il libre ? Même question avec $f_a(t) = \dfrac{1}{t^2+a^2}$.

EXERCICE 1-5.5 Soit $n \in \mathbf{N}$ tel que n ne soit le carré d'aucun entier.

1. Montrer que $\mathbf{Q}[\sqrt{n}] = \{a + b\sqrt{n},\ (a,b) \in \mathbf{Q}^2\}$ est un $\mathbf{Q}$-ev. En donner la dimension.

2. De même, on considère $E = \{a + b\sqrt{2} + c\sqrt{3},\ (a,b,c) \in \mathbf{Q}^3\}$. Montrer que c'est un $\mathbf{Q}$-ev et en donner la dimension. Déterminer le plus petit sous-corps (pour l'inclusion) de $\mathbf{C}$ contenant $\sqrt{2}$ et $\sqrt{3}$.

Chapitre 2

Applications linéaires

2-1 Généralités

2-1.1 Définition, exemples

DÉFINITION 2-1.1 *Si $\mathbb{E}$ et $\mathbb{F}$ sont deux espaces vectoriels sur le même corps commutatif $\mathbb{K}$, une application $u : \mathbb{E} \to \mathbb{F}$ est dite linéaire ssi elle vérifie :*

$$\begin{array}{ll} \forall x, y \in \mathbb{E} & u(x+y) = u(x) + u(y) \\ \forall x \in \mathbb{E} \quad \forall \lambda \in \mathbb{K} & u(\lambda x) = \lambda u(x) \end{array}$$

On dit alors de manière équivalente que u est un morphisme (d'espaces vectoriels).

Il est parfois nécessaire de préciser les structures d'espaces vectoriels intervenant. Par exemple, l'application $\mathbb{C} \to \mathbb{C}$ définie par $z \mapsto \overline{z}$ est $\mathbb{R}$-linéaire, mais elle n'est pas $\mathbb{C}$-linéaire.

Lorsque u est de plus bijective, on dit que c'est un **isomorphisme** d'espaces vectoriels. Lorsque $\mathbb{E} = \mathbb{F}$, on parle d'**endomorphisme** de l'espace vectoriel $\mathbb{E}$. Un endomorphisme bijectif de $\mathbb{E}$ est appelé **automorphisme de** $\mathbb{E}$.

On note $\mathcal{L}(\mathbb{E}, \mathbb{F})$ l'ensemble des morphismes de $\mathbb{E}$ vers $\mathbb{F}$ et $\mathcal{L}(\mathbb{E})$ l'ensemble des endomorphismes de $\mathbb{E}$. On notera ces ensembles respectivement $\mathcal{L}_{\mathbb{K}}(\mathbb{E}, \mathbb{F})$ et $\mathcal{L}_{\mathbb{K}}(\mathbb{E})$ s'il risque d'y avoir ambiguité sur le corps de base.

EXEMPLE 2-1.2 Si $\mathbb{E} = \mathbb{F}_1 \oplus \mathbb{F}_2$ est décomposé en somme de deux sous-espaces supplémentaires, le projecteur sur $\mathbb{F}_1$ parallèlement à $\mathbb{F}_2$ est un endomorphisme de $\mathbb{E}$.

EXEMPLE 2-1.3 Si $\mathbb{E}$ possède une base $(u_i)_{i\in I}$ et si $\mathbb{K}^{(I)}$ est l'espace vectoriel des familles de scalaires $(\alpha_i)_{i\in I}$ presque tous nuls, l'application

$$\mathbb{K}^{(I)} \to \mathbb{E} \qquad (\alpha_i)_{i\in I} \mapsto \sum_{i\in I} \alpha_i u_i$$

est un isomorphisme d'espaces vectoriels. En particulier si $\mathbb{E}$ est de dimension finie n, le choix d'une base de $\mathbb{E}$ permet de construire un isomorphisme $\mathbb{K}^n \to \mathbb{E}$.

PROPOSITION 2-1.4 Si $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$ et si $\mathbb{E}_1$ est un sous-espace vectoriel, la restriction $u|_{\mathbb{E}_1}$ de u à $\mathbb{E}_1$ définit un élément de $\mathcal{L}(\mathbb{E}_1,\mathbb{F})$.

Attention ! Si $u \in \mathcal{L}(\mathbb{E})$, on aura $u|_{\mathbb{E}_1} \in \mathcal{L}(\mathbb{E}_1,\mathbb{E})$. On ne pourra considérer qu'il s'agit d'un endomorphisme de $\mathbb{E}_1$ que **dans le cas très particulier où** $u(\mathbb{E}_1) \subset \mathbb{E}_1$, donc dans le cas où le sous-espace $\mathbb{E}_1$ est **stable par** u.

2-1.2 Espaces isomorphes

THÉORÈME 2-1.5 Si $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$ et $v \in \mathcal{L}(\mathbb{F},\mathbb{G})$, le composé $v \circ u$ est un morphisme de $\mathbb{E}$ vers $\mathbb{G}$. Si u est un isomorphisme $\mathbb{E} \to \mathbb{F}$, la bijection réciproque $u^{-1} : \mathbb{F} \to \mathbb{E}$ est un isomorphisme. Le composé de deux isomorphismes d'espaces vectoriels est encore un isomorphisme.

COROLLAIRE 2-1.6 L'ensemble des automorphismes d'un espace vectoriel $\mathbb{E}$ est un groupe pour la composition des applications.

Démonstration : C'est évidemment, d'après ce qui précède, un sous-groupe du groupe $(\Sigma_{\mathbb{E}}, \circ)$ des permutations de l'ensemble $\mathbb{E}$. L'élément neutre est $id_{\mathbb{E}}$. ■

DÉFINITION 2-1.7 *Deux espaces vectoriels $\mathbb{E}$ et $\mathbb{F}$ sont dits isomorphes s'il existe (au moins) un isomorphisme $u : \mathbb{E} \to \mathbb{F}$.*

Cette relation est évidemment réflexive, symétrique et transitive (théorème **2-1**.5). Il est clair que si $u : \mathbb{E} \to \mathbb{F}$ est un isomorphisme, tous les isomorphismes $u' : \mathbb{E} \to \mathbb{F}$ peuvent s'écrire $u' = u \circ v$ où v est un automorphisme de $\mathbb{E}$, ou encore $w \circ u$ où w est un automorphisme de $\mathbb{F}$.

EXEMPLE 2-1.8 Si $(\mathbb{F}_i)_{1\leqslant i\leqslant p}$ est une famille finie de sev d'un espace $\mathbb{E}$, dire que ces sous-espaces sont indépendants équivaut à affirmer que

$$\prod_{i=1}^{p} \mathbb{F}_i \mapsto \sum_{i=1}^{p} \mathbb{F}_i \qquad (x_1, \ldots, x_p) \mapsto \sum_{i=1}^{p} x_i$$

est un isomorphisme.

EXEMPLE 2-1.9 Si $\mathbb{S}_1$ et $\mathbb{S}_2$ sont deux supplémentaires d'un sous-espace $\mathbb{E}_1$ de $\mathbb{E}$, $\mathbb{S}_1$ et $\mathbb{S}_2$ sont isomorphes. En dimension finie, on pourrait utiliser le fait que $\mathbb{S}_1$ et $\mathbb{S}_2$ ont même dimension. Mais, si on réfléchit un peu à cet argument, on s'aperçoit qu'il n'est pas totalement satisfaisant : l'isomorphisme que l'on

considère dépend du choix de bases de $\mathbb{S}_1$ et $\mathbb{S}_2$ (voir le corollaire **2-1**.23) et n'est donc pas "intrinsèque", pas "géométrique". La solution suivante est préférable, et valable sans hypothèse de dimension : pour "envoyer" de manière bijective (et linéaire!) $\mathbb{S}_1$ sur $\mathbb{S}_2$, il est naturel de considérer la restriction à $\mathbb{S}_1$ du projecteur p sur $\mathbb{S}_2$ parallèlement à $\mathbb{E}_1$. On peut, par un petit abus de langage, considérer que l'espace d'arrivée est $\mathbb{S}_2$ et on définit ainsi un morphisme

$$u = p|_{\mathbb{S}_1} : \mathbb{S}_1 \to \mathbb{S}_2$$

Ce morphisme est injectif : si x et $y \in \mathbb{S}_1$ vérifient $u(x) = u(y)$, on a alors $p(x-y) = 0_{\mathbb{E}}$, ce qui signifie $x - y \in \mathbb{E}_1 \cap \mathbb{S}_1$ et donc $x = y$ puisque $\mathbb{E}_1$ et $\mathbb{S}_1$ sont supplémentaires. Vérifions enfin que u est surjectif : si z est un vecteur quelconque de $\mathbb{S}_2$, il se décompose de manière unique sous la forme $z = x + y$ dans la somme directe $\mathbb{E} = \mathbb{S}_1 \oplus \mathbb{E}_1$. On a alors

$$z = p(z) = p(x+y) = p(x) = u(x)$$

ce qui prouve bien la surjectivité de u.

EXEMPLE 2-1.10 Si $\mathbb{K}$ est un **corps fini,** sa **caractéristique** (c'est-à-dire le plus petit entier $p > 0$ vérifiant $p \cdot 1 = 0$) est un **nombre premier.** On montre alors (voir chapitre **17**) que $\mathbb{K}_1 = \{0, 1, 2.1, \ldots, (p-1).1\}$ est un sous-corps de $\mathbb{K}$ (isomorphe à $\mathbb{Z}/p\mathbb{Z}$), de cardinal p. $\mathbb{K}$ peut être considéré comme un $\mathbb{K}_1$-espace vectoriel. Il est évidemment de dimension finie, et si n est cette dimension, $\mathbb{K}$ est isomorphe (en tant qu'espace vectoriel) à $\mathbb{K}_1^n$ et est donc de cardinal p^n. **Le cardinal d'un corps fini est donc une puissance de sa caractéristique.** Il n'existe donc pas de corps fini de cardinal 24 par exemple. On démontre par contre (c'est un peu plus difficile) que si p est un nombre premier et n un entier non nul, il existe effectivement un corps fini de cardinal p^n, et que deux corps ayant ce cardinal sont isomorphes.

2-1.3 Définition d'un morphisme

Lorsqu'on connaît une base de l'espace de départ, le théorème suivant est un moyen commode de définir ou de caractériser un morphisme d'espaces vectoriels.

THÉORÈME 2-1.11 Si $\mathcal{B} = (e_i)_{i \in I}$ est une base de $\mathbb{E}$ et $(y_i)_{i \in I}$ est une famille quelconque de vecteurs de $\mathbb{F}$, il existe un unique morphisme $u : \mathbb{E} \to \mathbb{F}$ vérifiant

$$\forall i \in I \qquad u(e_i) = y_i$$

Un morphisme est donc entièrement caractérisé par la donnée des images des vecteurs d'une base de l'espace de départ.

Démonstration : Si $u : \mathbb{E} \to \mathbb{F}$ répond à la question et si x est un vecteur de $\mathbb{E}$ qui admet l'écriture

$$x = \sum_{i \in I} \alpha_i e_i$$

dans la base $\mathcal{B}$ (les α_i sont presque tous nuls), on a nécessairement

$$u(x) = \sum_{i \in I} \alpha_i u(e_i) = \sum_{i \in I} \alpha_i y_i$$

Réciproquement, on vérifie aisément que cette "formule" permet de définir sans ambiguïté une application $\mathbb{E} \to \mathbb{F}$, qui est linéaire et répond à la question. ■

EXERCICE 2-1.12 On note $(e_1, \ldots, e_p)$ la base canonique de $\mathbb{K}^p$. Si $\mathcal{F} = (x_1, \ldots, x_p)$ est une famille quelconque d'un espace vectoriel $\mathbb{E}$, il existe, d'après le théorème précédent, un unique morphisme

$$v : \mathbb{K}^p \to \mathbb{E} \quad \text{avec} \quad v(e_i) = x_i$$

Montrer que v est injectif ssi $\mathcal{F}$ est libre, surjectif ssi $\mathcal{F}$ est une famille génératrice de $\mathbb{E}$. Quelle est la condition pour que v soit un isomorphisme ?

2-1.4 "Recollement" linéaire d'applications linéaires

Lorsque l'espace $\mathbb{E}$ de départ est décomposé en somme directe de droites vectorielles (ce qui correspond en fait à la donnée d'une base de $\mathbb{E}$), la section précédente nous a montré comment caractériser un morphisme $u : \mathbb{E} \to \mathbb{F}$. Nous généralisons ce résultat à une décomposition $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_p$. Il suffit en fait d'étudier la cas où $\mathbb{E}$ est décomposé en somme directe de deux sous-espaces supplémentaires.

Si $\mathbb{E} = \mathbb{E}_1 \oplus \mathbb{E}_2$ et $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$, u est entièrement déterminé lorsqu'on connaît ses restrictions $u_1 = u|_{\mathbb{E}_1} \in \mathcal{L}(\mathbb{E}_1, \mathbb{F})$ et $u_2 = u|_{\mathbb{E}_2} \in \mathcal{L}(\mathbb{E}_2, \mathbb{F})$, puisque si $x \in \mathbb{E}$ se décompose dans la somme directe sous la forme $x = x_1 + x_2$, on a nécessairement

$$u(x) = u(x_1) + u(x_2) = u_1(x_1) + u_2(x_2)$$

On a la réciproque :

THÉORÈME 2-1.13 Si $\mathbb{E} = \mathbb{E}_1 \oplus \mathbb{E}_2$ et si l'on se donne deux morphismes $v_1 \in \mathcal{L}(\mathbb{E}_1, \mathbb{F})$ et $v_2 \in \mathcal{L}(\mathbb{E}_2, \mathbb{F})$, il existe un unique morphisme $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ tel que $v_1 = u|_{\mathbb{E}_1}$ et $v_2 = u|_{\mathbb{E}_2}$.

Démonstration : On raisonne par analyse-synthèse : si u existe et si $x = x_1 + x_2$ est comme précédemment, on a forcément

$$u(x) = v_1(x_1) + v_2(x_2)$$

On vérifie que réciproquement, cette "formule" définit bien un morphisme u de $\mathbb{E}$ vers $\mathbb{F}$ qui répond bien à la question. ■

On dit que u est obtenu en recollant linéairement les morphismes v_1 et v_2, et il nous arrivera de noter $u = v_1 \oplus v_2$ (bien que cette notation ne soit pas standard). Par exemple, le projecteur sur $\mathbb{E}_1$ parallèlement à $\mathbb{E}_2$ peut être représenté par $id_{\mathbb{E}_1} \oplus 0$ (avec un petit abus d'écriture, puisqu'on considère ici $id_{\mathbb{E}_1}$ comme un morphisme de $\mathbb{E}_1$ dans $\mathbb{E}$. 0 représente ici le morphisme nul $\mathbb{E}_2 \to \mathbb{E}$).

Ce théorème **2-1**.13 est peut-être à l'origine de la confusion (pour les débutants !) entre supplémentaire et complémentaire. Pour définir le morphisme u sur le vecteur $x \in \mathbb{E}$, on étudie d'abord le cas où $x \in \mathbb{E}_1$ puis celui où $x \in \mathbb{E}_2$. Si on connaît u sur $\mathbb{E}_1 \cup \mathbb{E}_2$, **en raisonnant par linéarité (en "recollant" les morphismes) on connaît u sur $\mathbb{E}_1 \oplus \mathbb{E}_2$**, bien que $\mathbb{E}_1 \cup \mathbb{E}_2 \neq \mathbb{E}_1 \oplus \mathbb{E}_2$.

La construction précédente se généralise évidemment au cas où $\mathbb{E}$ se décompose en $\mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_p$. On pourra ainsi recoller (sous-entendu "linéairement") des morphismes $v_i \in \mathcal{L}(\mathbb{E}_i, \mathbb{F})$ en un morphisme $u = v_1 \oplus \cdots \oplus v_p \in \mathcal{L}(\mathbb{E}, \mathbb{F})$

2-1.5 Images directes et réciproques

La linéarité d'une application $u : \mathbb{E} \to \mathbb{F}$ entraîne évidemment que l'image d'une combinaison linéaire (à coefficients presque tous nuls) d'une famille de $\mathbb{E}$ est la combinaison linéaire des images :

$$u\left(\sum_{i\in I} \lambda_i x_i\right) = \sum_{i\in I} \lambda_i u(x_i)$$

Comme on a de plus

$$u(0_{\mathbb{E}}) = 0_{\mathbb{F}}$$

(u est déjà un morphisme de groupes additifs), l'image d'un système lié est évidemment un système lié, et **une application linéaire "transporte" les relations de liaison.** L'objet principal de cette section est de voir à quelle condition on est assuré que l'image d'un système libre est libre. De manière équivalente, à quelle condition une application linéaire ne crée-t-elle pas de relations de liaison ?

2-1.5.1 Images de sous-espaces vectoriels

THÉORÈME 2-1.14 Si $u : \mathbb{E} \to \mathbb{F}$ est un morphisme, l'image directe par u d'un sous-espace de $\mathbb{E}$ est un sous-espace de $\mathbb{F}$, l'image réciproque d'un sous-espace de $\mathbb{F}$ est un sous-espace de $\mathbb{E}$.

Démonstration : Prouvons le pour les images réciproques. Si $\mathbb{F}_1$ est un sous-espace vectoriel de $\mathbb{F}$

$$\mathbb{E}_1 = u^{-1}(\mathbb{F}_1) = \{x \in \mathbb{E} \mid u(x) \in \mathbb{F}_1\}$$

n'est pas vide puisque $u(0_{\mathbb{E}}) = 0_{\mathbb{F}} \in \mathbb{F}_1$. Si x et $y \in \mathbb{E}_1$, et $\alpha \in \mathbb{K}$ on a

$$u(x + \alpha y) = u(x) + \alpha u(y)$$

Ce vecteur est combinaison linéaire de deux éléments de $\mathbb{F}_1$ et est donc dans $\mathbb{F}_1$. On a donc bien $x + \alpha y \in \mathbb{E}_1$. ■

DÉFINITION 2-1.15 (Noyau, image) *Si $u : \mathbb{E} \to \mathbb{F}$ est un morphisme, on appelle noyau de u, et on note* $\ker u$ *le sous-espace de* $\mathbb{E}$

$$\ker u = u^{-1}(0_{\mathbb{F}}) = \{x \in \mathbb{E} \mid u(x) = 0_{\mathbb{F}}\}$$

L'image de u est le sous-espace de $\mathbb{F}$

$$\operatorname{Im} u = u(\mathbb{E}) = \{y \in \mathbb{F} \mid \exists x \in \mathbb{E} \quad u(x) = y\}$$

THÉORÈME 2-1.16 Si $u : \mathbb{E} \to \mathbb{F}$ est un morphisme, u est injectif ssi

$$\ker u = \{0_{\mathbb{E}}\}$$

La surjectivité de u équivaut évidemment à l'égalité $\operatorname{Im} u = \mathbb{F}$.

Démonstration : Si u est injectif, $0_{\mathbb{F}}$ ne possède que l'antécédent $0_{\mathbb{E}}$. Réciproquement, si $\ker u = \{0_{\mathbb{E}}\}$ et $x, y \in \mathbb{E}$

$$u(x) = u(y) \Leftrightarrow u(x-y) = 0_{\mathbb{F}} \Leftrightarrow x - y \in \ker u$$

montre l'injectivité de u. ■

Attention ! Pour que cette caractérisation de l'injectivité soit utilisable, il faut avoir au préalable vérifié la linéarité de u.

EXERCICE 2-1.17 Montrer qu'un morphisme $u : \mathbb{E} \to \mathbb{F}$ permet de définir une bijection $\mathbb{S} \mapsto u(\mathbb{S})$ entre l'ensemble des sev de **E contenant** $\ker u$ et l'ensemble des sev de $\operatorname{Im} u$. Quelle est la bijection réciproque ?

2-1.5.2 Image directe d'une famille

Commençons par la définition de l'image d'une famille :

DÉFINITION 2-1.18 *Si f est une application d'un ensemble $\mathbb{X}$ vers un ensemble $\mathbb{Y}$, et si $(x_i)_{i\in I}$ est une famille d'éléments de $\mathbb{X}$, la famille $(f(x_i))_{i\in I}$ d'éléments de $\mathbb{Y}$, indexée elle aussi par I, est appelée image de cette famille par l'application f*

On a déjà vu que, par un morphisme, l'image directe d'une famille liée est une famille liée, et qu'un morphisme transporte les relations de liaison.

THÉORÈME 2-1.19 Un morphisme $u : \mathbb{E} \to \mathbb{F}$ est injectif ssi l'image de tout système libre de $\mathbb{E}$ est un système libre de $\mathbb{F}$. En d'autres termes, les morphismes injectifs sont exactement ceux qui ne créent pas de relations de liaison.

Démonstration : Si $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ est injectif et si $(x_i)_{i\in I}$ est une famille libre de vecteurs de $\mathbb{E}$, on a :

$$\sum_{i\in I} \lambda_i u(x_i) = 0_{\mathbb{F}} \Leftrightarrow u\left(\sum_{i\in I} \lambda_i x_i\right) = 0_{\mathbb{F}} \Leftrightarrow \sum_{i\in I} \lambda_i x_i = 0_{\mathbb{E}}$$

car le noyau de u est réduit à $\{0_{\mathbb{E}}\}$. Compte tenu de l'indépendance linéaire des x_i, tous les coefficients sont nuls, ce qui prouve l'indépendance des $(u(x_i))_{i\in I}$. Réciproquement, si tout système libre a une image libre, un vecteur non nul a pour image un vecteur non nul, ce qui prouve que $\ker u = \{0_{\mathbb{E}}\}$. ■

COROLLAIRE 2-1.20 Une application linéaire est injective si et seulement si elle conserve le rang de tout système de vecteurs.

PROPOSITION 2-1.21 Si $\mathcal{F}$ est une famille génératrice de $\mathbb{E}$ et $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$, $u(\mathcal{F})$ est une famille génératrice de $\operatorname{Im} u$. En particulier, pour vérifier que u est surjective, il suffit de vérifier que l'image d'une famille génératrice de $\mathbb{E}$ engendre $\mathbb{F}$.

THÉORÈME 2-1.22 $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$ est un isomorphisme d'espaces vectoriels si et seulement si l'image d'une base de E est une base de $\mathbb{F}$.

Démonstration : Soit $\mathcal{B} = (e_i)_{i \in I}$ est une base de $\mathbb{E}$. Supposons d'abord que $u(\mathcal{B})$ soit une base de $\mathbb{F}$. D'après la proposition précédente, u est surjective. De plus, si $x = \sum_{i \in I} \alpha_i e_i$ est un vecteur de $\ker u$, on a

$$\sum_{i \in I} \alpha_i u(e_i) = 0_{\mathbb{F}}$$

ce qui entraîne, par indépendance de $u(\mathcal{B})$,

$$\forall i \in I \quad \alpha_i = 0 \quad \text{et donc} \quad x = 0_{\mathbb{E}}$$

ce qui prouve l'injectivité de u. Réciproquement, si u est un isomorphisme, une base $\mathcal{B}$ est libre donc (u injective) $u(\mathcal{B})$ est libre. $\mathcal{B}$ est génératrice de $\mathbb{E}$ donc (u surjective) $u(\mathcal{B})$ engendre $\mathbb{F}$. $u(\mathcal{B})$ est donc libre et génératrice de $\mathbb{F}$, c'est une base de $\mathbb{F}$. ■

COROLLAIRE 2-1.23 Si $\mathbb{E}$ est un $\mathbb{K}$-ev de dimension finie n, un $\mathbb{K}$-ev $\mathbb{F}$ est isomorphe à $\mathbb{E}$ si et seulement s'il est également de dimension n.

Démonstration : La condition est évidemment nécessaire puisqu'un isomorphisme $u : \mathbb{E} \to \mathbb{F}$ transforme une base de $\mathbb{E}$ en une base de $\mathbb{F}$. Elle est suffisante, puisque l'existence d'une base de cardinal n dans $\mathbb{E}$ et $\mathbb{F}$ permet, en utilisant le théorème **2-1**.11, de construire un isomorphisme entre $\mathbb{E}$ et $\mathbb{F}$. ■

2-1.6 Codimension d'un sous-espace. Hyperplan

L'exemple **2-1**.9 permet de donner la définition suivante :

DÉFINITION 2-1.24 *Si un sous-espace $\mathbb{E}_1$ d'un espace $\mathbb{E}$ possède un supplémentaire de dimension finie, tous ses supplémentaires ont la même dimension. Cette dimension sera appelée codimension du sous-espace $\mathbb{E}_1$ (dans $\mathbb{E}$) et sera notée $\operatorname{codim} \mathbb{E}_1$. On dira que $\mathbb{E}_1$ est de codimension infinie si $\mathbb{E}_1$ ne possède pas de supplémentaire de dimension finie.*

DÉFINITION 2-1.25 *Un sous-espace $\mathbb{H}$ de $\mathbb{E}$ est dit hyperplan de $\mathbb{E}$ ssi $\mathbb{H}$ est de codimension 1, ce qui signifie que $\mathbb{H}$ possède un supplémentaire qui est une droite vectorielle.*

Nous reviendrons sur ces définitions dans le chapitre sur la dualité. En attendant, quelques exercices :

EXERCICE 2-1.26 Si $\mathbb{H}$ est un hyperplan de $\mathbb{E}$ et si a est un vecteur de $\mathbb{E}$ qui n'est pas dans $\mathbb{H}$ montrer que

$$\mathbb{H} \oplus \text{vect}\,(a) = \mathbb{E}$$

EXERCICE 2-1.27 Si $\mathbb{H}$ est un hyperplan de $\mathbb{E}$ et $\mathbb{F}$ un sev de $\mathbb{E}$, montrer que

$$\mathbb{F} \cap \mathbb{H} = \mathbb{F} \quad \text{ou} \quad \mathbb{F} \cap \mathbb{H} \text{ est un hyperplan de } \mathbb{F}$$

EXERCICE 2-1.28 Montrer qu'un sous-espace $\mathbb{E}_1$ de $\mathbb{E}$ est de codimension finie ssi l'ensemble des dimensions des sous-espaces (de dimension finie) de $\mathbb{E}$ indépendants de $\mathbb{E}_1$ est majoré. Montrer que ceci équivaut aussi à dire que tout sous-espace indépendant de $\mathbb{E}_1$ est de dimension finie.

EXERCICE 2-1.29 Déduire de l'exercice précédent que, si $\mathbb{F}$ et $\mathbb{G}$ sont deux sous-espaces de codimension finie dans $\mathbb{E}$, il en est de même de $\mathbb{F} + \mathbb{G}$ et $\mathbb{F} \cap \mathbb{G}$ et

$$\text{codim}\,(\mathbb{F}) + \text{codim}\,(\mathbb{G}) = \text{codim}\,(\mathbb{F} + \mathbb{G}) + \text{codim}\,(\mathbb{F} \cap \mathbb{G})$$

2-1.7 Equation linéaire. Sous-espace affine

DÉFINITION 2-1.30 *On appelle équation linéaire toute équation de la forme*

$$a\,(x) = b$$

où a est une application linéaire d'un espace vectoriel $\mathbb{E}$ vers un espace $\mathbb{F}$, b est un vecteur donné de $\mathbb{F}$ (on dit souvent que b est le second membre de l'équation) et x est un vecteur inconnu dans $\mathbb{F}$.

DÉFINITION 2-1.31 *L'équation homogène associée à l'équation précédente (on dit aussi équation sans second membre) est l'équation*

$$a(x) = 0_{\mathbb{F}}$$

L'ensemble de ses solutions est le sous-espace de $\mathbb{E}$ égal à $\ker a$. Le vecteur nul $0_{\mathbb{E}}$ est en particulier toujours solution de l'équation homogène.

La résolution théorique d'une équation linéaire est fort simple :

- Si $b \notin \operatorname{Im} a$ (ce qui n'est évidemment possible que lorsque a n'est pas surjectif) l'équation n'a pas de solution.
- Sinon, si on connaît une solution particulière x_0, on a :

  $$a\,(x) = b = a\,(x_0) \Leftrightarrow a\,(x - x_0) = 0_{\mathbb{F}} \Leftrightarrow x - x_0 \in \ker a$$

  et l'ensemble des solutions est alors

  $$S = x_0 + \ker a = \{x_0 + y\ ,\ y \in \ker a\}$$

  On dit alors que la solution générale de l'équation avec second membre est somme d'une solution particulière de cette équation et de la solution générale de l'équation homogène associée. En particulier, la différence de deux solutions de l'équation avec second membre est solution de l'équation homogène.

On notera que la **linéarité de** a est fondamentale dans ce qui précède. Remarquons que cette propriété de linéarité est une traduction du "principe de superposition des solutions" : si x_0 est solution de $a(x) = b_0$ et x_1 est solution de $a(x) = b_1$, $x_0 + \alpha x_1$ est solution de $a(x) = b_0 + \alpha b_1$.

DÉFINITION 2-1.32 *Si* $\mathbb{D}$ *est un sous-espace vectoriel de* $\mathbb{E}$ *et* x_0 *est un vecteur quelconque de* $\mathbf{E}$*, on appelle sous-espace affine de* $\mathbf{E}$ *passant par* x_0 *et de direction* $\mathbb{D}$ *le sous-ensemble de* $\mathbb{E}$

$$x_0 + \mathbb{D} = \{x_0 + d, \quad d \in \mathbb{D}\}$$

Il est alors facile de voir que, pour tout $x \in x_0 + \mathbb{D}$ on a aussi

$$x + \mathbb{D} = x_0 + \mathbb{D}$$

EXERCICE 2-1.33 Montrer qu'une partie $\mathbf{A}$ de $\mathbf{E}$ est un sous-espace affine si et seulement si

$$\{x - y \mid x \in \mathbf{A} \text{ et } y \in \mathbf{A}\}$$

est un sous-espace vectoriel de $\mathbf{E}$, et que ce sous-espace est alors exactement la direction de $\mathbf{A}$.

Un sous-espace affine est donc caractérisé par la donnée d'un élément quelconque lui appartenant et par sa direction, qui est un sous-espace vectoriel de $\mathbb{E}$. Un sous-espace vectoriel est un sous-espace affine passant par "l'origine" $0_{\mathbb{E}}$. Avec ce vocabulaire, l'ensemble des solutions d'une équation linéaire $a(x) = b$ est soit réduite à $\emptyset$, soit un sous-espace affine de l'espace de départ, de direction égale à $\ker a$.

EXERCICE 2-1.34 Montrer que deux sous-espaces affines $x + \mathbf{D}$ et $x' + \mathbb{D}'$ ont une intersection non vide ssi

$$x - x' \in \mathbf{D} + \mathbb{D}'$$

et que l'intersection est alors un sous-espace affine de direction $\mathbf{D} \cap \mathbb{D}'$. En particulier, deux sous-espaces affines dont les directions sont supplémentaires ont toujours une intersection réduite à un point.

EXERCICE 2-1.35 On dit que le sous-espace affine $x + \mathbf{D}$ est parallèle à $x' + \mathbb{D}'$ ssi $\mathbf{D}$ est inclus dans $\mathbb{D}'$. Montrer qu'on définit ainsi une relation réflexive et transitive dans l'ensemble des sous-espaces affines de $\mathbf{E}$ et que la relation symétrique associée (deux sous-espaces affines sont dits parallèles s'ils ont mêmes directions) est une relation d'équivalence.

Bien évidemment, on appelle **hyperplan affine** tout sous-espace affine dont la direction est un hyperplan vectoriel, notion sur laquelle nous reviendrons dans le chapitre suivant.

2-1.8 Théorème fondamental d'isomorphisme

Lorsqu'on fait opérer une application linéaire u non injective sur un espace vectoriel $\mathbb{E}$, on peut considérer qu'à l'arrivée, on a perdu de l'information sur les vecteurs de $\mathbb{E}$ puisque deux vecteurs de $\mathbb{E}$ différant d'un vecteur de $\ker u$ sont indiscernables si on ne regarde que les images. Le théorème suivant montre que, en se restreignant à un supplémentaire de $\ker u$ dans $\mathbb{E}$, il n'y a pas de perte d'information :

THÉORÈME 2-1.36 Si $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ et si $\mathbb{S}$ est un supplémentaire de $\ker u$ dans $\mathbb{E}$, la restriction de u à $\mathbb{S}$ induit un isomorphisme de $\mathbb{S}$ vers $\operatorname{Im} u$.

Démonstration : On a évidemment

$$\ker u|_{\mathbb{S}} = \ker u \cap \mathbb{S}$$

et donc, puisque $\ker u$ et $\mathbb{S}$ sont indépendants, $u|_{\mathbb{S}} : \mathbb{S} \to \operatorname{Im} u$ est injective. De plus, si y est un vecteur quelconque de $\operatorname{Im} u$, il admet au moins un antécédent x par u, qui se décompose dans la somme $\mathbb{E} = \mathbb{S} \oplus \ker u$ sous la forme $x = x_{\mathbb{S}} + x_k$, et on a alors

$$y = u(x) = u(x_{\mathbb{S}} + x_k) = u(x_{\mathbb{S}}) = u|_{\mathbb{S}}(x_{\mathbb{S}})$$

ce qui prouve la surjectivité de $u|_{\mathbb{S}} : \mathbb{S} \to \operatorname{Im} u$. ■

2-2 Calculs sur les morphismes

On a déja vu (théorème **2-1**.5) que, lorsqu'il est défini, le composé de deux morphismes d'espaces vectoriels est aussi une application linéaire. Le fait que les espaces d'arrivée sont des espaces vectoriels va permettre de définir d'autres opérations sur les morphismes, qui vont avoir l'avantage de bien se comporter par rapport à la composition.

2-2.1 Structure d'espace vectoriel de $\mathcal{L}_{\mathbb{K}}(\mathbb{E}, \mathbb{F})$

Si $\mathbb{E}$ et $\mathbb{F}$ sont deux $\mathbb{K}$-ev, rappelons (exemple **1-1**.5) que l'ensemble $\mathbb{F}^{\mathbb{E}}$ des applications de $\mathbb{E}$ dans $\mathbb{F}$ est muni d'une structure naturelle de $\mathbb{K}$-ev, les opérations sur les applications étant définies à partir des opérations sur les images.

THÉORÈME 2-2.1 Si $\mathbb{E}$ et $\mathbb{F}$ sont deux $\mathbb{K}$-ev, $\mathcal{L}_{\mathbb{K}}(\mathbb{E}, \mathbb{F})$ est un sous-espace vectoriel de $\mathbb{F}^{\mathbb{E}}$.

Démonstration : $\mathcal{L}_{\mathbb{K}}(\mathbb{E}, \mathbb{F})$ n'est pas vide, puisqu'il contient au moins l'application identiquement nulle de $\mathbb{E}$ dans $\mathbb{F}$ (qui est le vecteur nul de $\mathbb{F}^{\mathbb{E}}$). On vérifie ensuite très aisément que, si u et v sont des morphismes de $\mathbb{E}$ vers $\mathbb{F}$, il en est de même de $u+\alpha v$ pour α quelconque dans $\mathbb{K}$. ■

Avec cette structure, le théorème **2-1**.13 peut se ré-écrire :

THÉORÈME 2-2.2 Si $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_p$, l'application

$$\begin{aligned} \mathcal{L}_{\mathbb{K}}(\mathbb{E},\mathbb{F}) &\to \prod_{i=1}^{p} \mathcal{L}_{\mathbb{K}}(\mathbb{E}_i,\mathbb{F}) \\ u &\longmapsto \left(u|_{\mathbb{E}_1},\ldots,u|_{\mathbb{E}_p}\right) \end{aligned}$$

est un isomorphisme d'espaces vectoriels.

Considérer l'isomorphisme réciproque, c'est recoller linéairement des applications linéaires.

2-2.2 Propriétés de la composition des morphismes

La composition des applications est associative. Elle le reste bien évidemment lorsqu'on se restreint à travailler sur des morphismes d'espaces vectoriels. On a de plus les propriétés de distributivité et d'exportativité des scalaires :

THÉORÈME 2-2.3 La composition des morphismes est distributive à droite et à gauche par rapport à l'addition.

De manière plus précise, cela signifie que, si $\mathbb{E},\mathbb{F}$ et $\mathbb{G}$ sont trois espaces vectoriels

$$\begin{aligned} u \in \mathcal{L}(\mathbb{F},\mathbb{G}),\ v \text{ et } w \in \mathcal{L}(\mathbb{E},\mathbb{F}) &\Rightarrow u \circ (v+w) = u \circ v + u \circ w \\ u \in \mathcal{L}(\mathbb{E},\mathbb{F}),\ v \text{ et } w \in \mathcal{L}(\mathbb{F},\mathbb{G}) &\Rightarrow (v+w) \circ u = v \circ u + w \circ u \end{aligned}$$

La démonstration est évidente. Il est à noter que seule la première propriété utilise la linéarité.

THÉORÈME 2-2.4 Pour multiplier un composé de morphismes par un scalaire, il suffit de multiplier un des facteurs par ce scalaire.

Soit, pour $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$, $v \in \mathcal{L}(\mathbb{F},\mathbb{G})$ et $\alpha \in \mathbb{K}$

$$\alpha \cdot (v \circ u) = (\alpha v) \circ u = v \circ (\alpha u)$$

Ici encore, seule la deuxième égalité utilise une propriété de linéarité. On peut résumer les deux théorèmes précédents par l'énoncé suivant :

PROPOSITION 2-2.5 Si $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$, les applications

$$\begin{aligned} \mathcal{L}(\mathbb{F},\mathbb{G}) &\to \mathcal{L}(\mathbb{E},\mathbb{G}) \quad v \mapsto v \circ u \\ \mathcal{L}(\mathbb{G},\mathbb{E}) &\to \mathcal{L}(\mathbb{G},\mathbb{F}) \quad v \mapsto u \circ v \end{aligned}$$

sont linéaires.

Etant donné ce comportement sympathique de la composition des morphismes, la composée $v \circ u$ sera souvent appelée "produit" de v par u et sera simplement notée vu, lorsqu'il n'y a pas de confusion possible avec un autre produit qui serait défini entre les objets manipulés.

2-2.3 Structure d'algèbre de $\mathcal{L}_{\mathbb{K}}(\mathbb{E})$

Il résulte des théorèmes précédents que, si $\mathbb{E}$ est un $\mathbb{K}$-ev :

- $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +, \cdot)$ est un $\mathbb{K}$-espace vectoriel.
- $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +, \circ)$ possède une structure d'anneau, puisque $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +)$ est un groupe commutatif, que $\circ$ est associative, est distributive par rapport à l'addition et possède un élément neutre $id_{\mathbb{E}}$.
- Il existe une relation entre multiplication interne (c'est ici la composition) et multiplication externe. C'est l'exportativité des scalaires :

$$\forall u, v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}) \quad \forall \alpha \in \mathbb{K} \qquad \alpha(uv) = (\alpha u)v = u(\alpha v)$$

Pour ces trois raisons, on dit que $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +, ., \circ)$ est une $\mathbb{K}$-algèbre[1], conformément à la définition suivante :

DÉFINITION 2-2.6 *Une structure de $\mathbb{K}$-algèbre sur un ensemble $\mathbb{A}$ est la donnée de deux lois de composition internes (notées habituellement $+$ et $\times$) et d'une multiplication externe à domaine d'opérateurs égal à $\mathbb{K}$ telles que*

$$\begin{cases} \bullet\ (\mathbb{A}, +, \cdot) \text{ est un } \mathbb{K}\text{-ev} \\ \bullet\ (\mathbb{A}, +, \times) \text{ est un anneau, l'élément neutre pour } \times \text{ étant noté } 1_{\mathbb{A}} \\ \bullet\ \forall a, b \in \mathbb{A} \quad \forall \alpha \in \mathbb{K} \qquad \alpha(a \times b) = (\alpha a) \times b = a \times (\alpha b) \end{cases}$$

Par définition, les anneaux (donc aussi les algèbres) sont supposés posséder un élément neutre pour la multiplication. La définition donnée ci-dessus correspond à ce qu'on appelait autrefois une algèbre unitaire.

Les termes "algèbre commutative", "algèbre intègre" se rapportent aux propriétés de l'anneau $(\mathbb{A}, +, \times)$. Dans le premier cas, la multiplication est commutative, dans le second cas il n'y a pas de diviseurs de zéro ce qui signifie

$$\forall a, b \in \mathbb{A} \quad ab = 0 \Rightarrow a = 0 \text{ ou } b = 0$$

Ceci équivaut en fait à dire que tout élément non nul est régulier (on dit aussi simplifiable) à droite et à gauche, puisque

$$ax = ay \Leftrightarrow a(x - y) = 0$$

DÉFINITION 2-2.7 *Si $(\mathbb{A}, +, \times, .)$ et $(\mathbb{B}, +, \times, .)$ sont deux $\mathbb{K}$-algébres, une application $\varphi : \mathbb{A} \to \mathbb{B}$ est un morphisme d'algèbres si φ est à la fois un morphisme d'espaces vectoriels et un morphisme d'anneaux, c'est-à-dire*

$$\varphi(x + y) = \varphi(x) + \varphi(y), \quad \varphi(\alpha x) = \alpha\varphi(x)$$
$$\varphi(xy) = \varphi(x)\varphi(y) \quad \text{et} \quad \varphi(1_{\mathbb{A}}) = \varphi(1_{\mathbb{B}})$$

[1] Lorsque l'espace $\mathbb{E}$ est réduit à $\{0_{\mathbb{E}}\}$, l'espace $\mathcal{L}_{\mathbb{K}}(\mathbb{E})$ est réduit au seul morphisme nul. Si on respecte strictement les définitions du programme, on ne peut plus vraiment dire que c'est une algèbre, puisque l'élément neutre pour la multiplication dans un anneau est toujours supposé distinct de l'élément neutre pour l'addition.

On parlera d'isomorphisme d'algèbres si de plus φ est bijective.

Le noyau d'un morphisme d'algèbres $\varphi : \mathbb{A} \to \mathbb{B}$ est évidemment un sous-espace vectoriel de $\mathbb{A}$. Il possède une propriété plus forte que la stabilité pour la multiplication (voir chapitre sur les anneaux) :

$$\forall a \in \mathbb{A} \quad \forall x \in \ker\varphi \qquad ax \text{ et } xa \in \ker\varphi$$

DÉFINITION 2-2.8 *Une sous-algèbre d'une algèbre $(\mathbb{A}, +, \times, .)$ est une partie $\mathbb{A}_1 \subset \mathbb{A}$ qui est à la fois un sous-espace vectoriel et un sous-anneau, c'est-à-dire*

$$\forall x, y \in \mathbb{A}_1, \quad \forall \alpha \in \mathbb{K} \qquad x + y,\ \alpha x,\ xy \in \mathbb{A}_1 \qquad et \quad 1_{\mathbb{A}} \in \mathbb{A}_1$$

$(\mathbb{A}_1, +, \times, .)$ est alors évidemment elle-même une $\mathbb{K}$-algèbre.

EXEMPLE 2-2.9 Si $\mathbb{L}$ est un corps commutatif contenant $\mathbb{K}$, $\mathbb{L}$ peut être considéré comme une $\mathbb{K}$-algèbre, qui est alors évidemment commutative et intègre.

EXEMPLE 2-2.10 Un autre exemple (très important) d'algèbre (intègre et commutative) est donné par $(\mathbb{K}[X], +, \times, \cdot)$, algèbre des polynômes à une indéterminée sur le corps $\mathbb{K}$.

Si $\mathbb{E}$ est un $\mathbb{K}$-espace vectoriel de dimension 1, un endomorphisme de $\mathbb{E}$ est nécessairement une homothétie c'est-à-dire est de la forme $\lambda\, id_{\mathbb{E}}$. En effet, si e_1 est un vecteur non nul de $\mathbb{E}$, il forme à lui seul une base de $\mathbb{E}$. Si $u \in \mathcal{L}(\mathbb{E})$, il existe un scalaire α tel que $u(e_1) = \alpha e_1$, et par linéarité on a forcément $u = \alpha\, id_{\mathbb{E}}$ (*cf.* théorème **2-1**.11). En dimension 1, l'algèbre $\mathcal{L}(\mathbb{E})$ est isomorphe au corps $\mathbb{K}$, par la correspondance $\mathbb{K} \to \mathcal{L}(\mathbb{E})$, $\lambda \mapsto \lambda\, id_{\mathbb{E}}$. Elle est donc dans ce cas commutative et intègre.

Dès que $\mathbb{E}$ contient deux vecteurs indépendants, $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +, ., \circ)$ n'est plus intègre ni commutative :

Si e_1 et e_2 sont deux vecteurs indépendants, en admettant que vect(e_1, e_2) possède un supplémentaire $\mathbb{S}$ dans $\mathbb{E}$, les endomorphismes u et v caractérisés par

$$u(e_1) = e_1,\ u(e_2) = e_1 + e_2,\ v(e_1) = e_2 \text{ et } v(e_2) = e_1, \quad u|_{\mathbb{S}} = v|_{\mathbb{S}} = id_{\mathbb{S}}$$

(voir le théorème **2-1**.13) ne commutent pas puisque

$$uv(e_1) = e_1 + e_2 \quad \text{et} \quad vu(e_1) = e_2$$

De même, les projecteurs p sur vect(e_1) parallèlement à vect$(e_2) \oplus \mathbb{S}$ et q sur vect$(e_2) \oplus \mathbb{S}$ parallèlement à vect(e_1) sont non nuls et vérifient $pq = qp = 0_{\mathcal{L}(\mathbb{E})}$. Ils forment donc un couple de diviseurs de 0. Bien sûr dans ce cas, $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}), +, ., \circ)$ contient strictement l'ensemble $\mathbb{K} \cdot id_{\mathbb{E}}$ des homothéties, qui forme une sous-algèbre isomorphe à $\mathbb{K}$.

2-2.4 Exercice : centre de l'algèbre $\mathcal{L}(\mathbb{E})$

Si $(\mathbb{A}, +, \times)$ est un anneau, on appelle centre de $\mathbb{A}$ l'ensemble $\mathcal{C}(\mathbb{A})$ des éléments de $\mathbb{A}$ qui commutent avec tous les éléments de $\mathbb{A}$ (cette notion n'a évidemment d'intérêt que dans le cas où $\mathbb{A}$ n'est pas commutatif). On vérifie aisément que $\mathcal{C}(\mathbb{A})$ est un sous-anneau de $\mathbb{A}$. Si $\mathbb{A}$ a de plus une structure d'algèbre, $\mathcal{C}(\mathbb{A})$ en est une sous-algèbre.

On se propose de démontrer que, si $\mathbb{E}$ est un espace vectoriel, le centre de $\mathcal{L}(\mathbb{E})$ est exactement la sous-algèbre des homothéties $\mathbb{K} \cdot id_{\mathbb{E}}$. Nous reverrons ultérieurement une démonstration matricielle de ce résultat lorsque $\mathbb{E}$ est de dimension finie. On en propose ici une démonstration géométrique, en admettant que, dans $\mathbb{E}$, tout sous-espace possède au moins un supplémentaire.

- Vérifier d'abord que $\mathbb{K} \cdot id_{\mathbb{E}} \subset \mathcal{C}(\mathcal{L}(\mathbb{E}))$.
- Montrer ensuite que, si $u \in \mathcal{L}(\mathbb{E})$ est tel que

  $$\forall x \in \mathbb{E} \qquad \{x, u(x)\} \text{ est liée}$$

  alors u est une homothétie.
- Il suffit alors, pour conclure, de montrer que, si $u \in \mathcal{C}(\mathcal{L}(\mathbb{E}))$, toute droite vectorielle de $\mathbb{E}$ est stable par u. On pourra pour cela utiliser le lemme suivant, sur lequel nous reviendrons en détail dans le chapitre sur la réduction des endomorphismes :

LEMME 2-2.11 Si u et v sont deux endomorphismes de $\mathbb{E}$ qui commutent, le noyau (et l'image) de v sont stables par u.

2-2.5 Groupe linéaire $\mathcal{GL}(\mathbb{E})$

Nous avons déjà vu (corollaire **2-1**.6) que l'ensemble des automorphismes d'un espace vectoriel forme un groupe pour la composition des morphismes. C'est en fait le groupe des éléments inversibles (pour la multiplication) dans l'algèbre $\mathcal{L}(\mathbb{E})$ (on dit aussi parfois le groupe des unités de l'algèbre $\mathcal{L}(\mathbb{E})$). Ce groupe est appelé **groupe linéaire de l'espace vectoriel** $\mathbb{E}$ et on le note $(\mathcal{GL}(\mathbb{E}), \circ)$ ou $(\mathrm{Aut}(\mathbb{E}), \circ)$.

En dimension 1, ce groupe est isomorphe à $(\mathbb{K}^*, \times)$, puisqu'une homothétie est bijective ssi son rapport est non nul. Il est donc commutatif dans ce cas .

Par contre, dès que la dimension de l'espace est supérieure à 2, $(\mathcal{GL}(\mathbb{E}), \circ)$ n'est pas commutatif. Pour le prouver, on peut utiliser l'exemple donné pour prouver la non commutativité de $\mathcal{L}(\mathbb{E})$.

EXERCICE 2-2.12 Si $\mathbf{E} = \mathbf{E}_1 \oplus \mathbb{E}_2$, si $u_1 \in \mathcal{L}(\mathbb{E}_1)$ et $u_2 \in \mathcal{L}(\mathbf{E}_2)$, montrer que l'endomorphisme de $\mathbf{E}$ $u = u_1 \oplus u_2$ obtenu en recollant u_1 et u_2 est un automorphisme de $\mathbf{E}$ ssi $u_1 \in \mathcal{GL}(\mathbf{E}_1)$ et $u_2 \in \mathcal{GL}(\mathbb{E}_2)$. Montrer qu'on a alors

$$u^{-1} = u_1^{-1} \oplus u_2^{-1}$$

Montrer que l'application

$$\mathcal{GL}(\mathbf{E}_1) \times \mathcal{GL}(\mathbf{E}_1) \to \mathcal{GL}(\mathbf{E}_1) \qquad (u_1, u_2) \mapsto u_1 \oplus u_2$$

est un morphisme de groupes injectif. (Attention : ce n'est pas en général un isomorphisme. Un automorphisme quelconque de $\mathbb{E}$ ne laisse pas forcément $\mathbb{E}_1$ et $\mathbb{E}_2$ stables).

2-2.6 Projecteurs et involutions

2-2.6.1 Caractérisation des projecteurs

Nous avons vu à la section **1-3.8** la définition géométrique d'un projecteur. En particulier, si $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_m$ est une écriture de $\mathbb{E}$ comme somme directe de sous-espaces indépendants, on sait qu'on peut associer à cette décomposition une famille de projecteurs $(p_i)_{1 \leqslant i \leqslant m}$, où

$$p_i \text{ est le projecteur sur } \mathbb{E}_i \text{ parallèlement à } \bigoplus_{j \neq i} \mathbb{E}_j$$

On a dans l'algèbre $\mathcal{L}(\mathbb{E})$

$$p_i^2 = p_i \qquad p_i p_j = 0 \text{ si } i \neq j \qquad \text{et} \qquad \sum_{i=1}^{m} p_i = id_{\mathbb{E}}$$

la dernière égalité signifiant simplement que l'on reconstitue tout vecteur x de $\mathbb{E}$ à partir de ses projections $x_i = p_i(x)$ sur les sous-espaces $\mathbb{E}_i$ par $x = x_1 + \cdots + x_m$.

Il est remarquable de constater que, dans l'algèbre $\mathcal{L}(\mathbb{E})$, les projecteurs sont caractérisés par une condition algébrique très simple : ce sont les solutions de l'équation

$$X^2 = X$$

THÉORÈME 2-2.13 Soit u un endomorphisme d'un espace vectoriel $\mathbb{E}$. u est un projecteur si et seulement si $u^2 = u$. Lorsque cette égalité est vérifiée, $\ker u$ et $\operatorname{Im} u$ sont deux sous-espaces supplémentaires de $\mathbb{E}$ et u est le projecteur sur $\operatorname{Im} u$ parallèlement à $\ker u$.

Démonstration : La condition $u^2 = u$ est évidemment nécessaire pour que u soit un projecteur. Réciproquement, si $u^2 = u$, montrons que $\ker u$ et $\operatorname{Im} u$ sont deux sous-espaces supplémentaires de $\mathbb{E}$. Si $x \in \ker u \cap \operatorname{Im} u$, il existe $y \in \mathbb{E}$ avec $x = u(y)$. On a

$$u(x) = 0_{\mathbb{E}} = u^2(y) = u(y) = x$$

Ces deux sev sont indépendants. De plus, tout vecteur de $\mathbb{E}$ peut s'écrire

$$x = u(x) + (x - u(x))$$

qui est une décomposition de x dans $\operatorname{Im} u \oplus \ker u$. Les deux espaces sont supplémentaires, et la composante de x selon $\operatorname{Im} u$ est égale à $u(x)$. u est donc bien le projecteur sur $\operatorname{Im} u$ parallèlement à $\ker u$. ■

EXERCICE 2-2.14 Si p est un projecteur d'un espace vectoriel $\mathbb{E}$, donner une condition nécessaire et suffisante pour qu'un endomorphisme v de $\mathbb{E}$ commute avec p.

2-2.6.2 Symétries et involutions

On suppose, dans cette section, que $\mathbb{K}$ est un corps **de caractéristique différente de 2**. Dans ce cas, si x est un vecteur non nul de $\mathbb{E}$, les vecteurs x et $-x$ sont distincts.

DÉFINITION 2-2.15 *Si $\mathbb{E} = \mathbb{E}_1 \oplus \mathbb{E}_2$, on appelle symétrie par rapport à $\mathbb{E}_1$ parallèlement à $\mathbb{E}_2$ l'endomorphisme s de $\mathbb{E}$ défini par*

$$s = id_{\mathbb{E}_1} \oplus (-id_{\mathbb{E}_2})$$

Il s'agit donc du morphisme qui, à un vecteur décomposé dans $\mathbb{E}_1 \oplus \mathbb{E}_2$

$$x = x_1 + x_2$$

associe le vecteur

$$s(x) = x_1 - x_2$$

Il s'agit clairement d'un automorphisme de $\mathbb{E}$ qui est **involutif**, c'est-à-dire vérifie $s \circ s = id_{\mathbb{E}}$, soit encore $s = s^{-1}$. Dans $\mathcal{L}(\mathbb{E})$, s s'exprime à l'aide du projecteur p sur $\mathbb{E}_1$ parallèlement à $\mathbb{E}_2$:

$$s = 2p - id_{\mathbb{E}}$$

Ici encore, les symétries, définies géométriquement, sont caractérisées par l'égalité $s^2 = id_{\mathbb{E}}$. Ce sont les solutions de l'équation

$$X^2 = 1$$

dans l'algèbre $\mathcal{L}(\mathbb{E})$.

THÉORÈME 2-2.16 Un endomorphisme u d'un espace vectoriel $\mathbb{E}$ sur un corps de caractéristique différente de 2 est une symétrie ssi $u^2 = id_{\mathbb{E}}$. Lorsque cette équation est vérifiée, $\mathbb{E}$ est somme directe des sous-espaces

$$\mathbb{E}_1 = \ker(u - id_{\mathbb{E}}) \text{ et } \mathbb{E}_2 = \ker(u + id_{\mathbb{E}})$$

u est alors la symétrie par rapport à $\mathbb{E}_1$ parallèlement à $\mathbb{E}_2$.

Démonstration : Si $u^2 = id_{\mathbb{E}}$, on essaie d'écrire comme précédemment

$$u = 2p - id_{\mathbb{E}}$$

avec p projecteur. Comme la caractéristique de $\mathbb{K}$ est différente de 2, on pose

$$p = \frac{1}{2}(u + id_{\mathbb{E}})$$

On vérifie que p est un projecteur :

$$p^2 = \frac{1}{4}(u + id_{\mathbb{E}})^2 = \frac{1}{4}(u^2 + 2u + id_{\mathbb{E}}) = \frac{1}{2}(u + id_{\mathbb{E}}) = p$$

puisque $u^2 = id_{\mathbb{E}}$. u est donc bien la symétrie par rapport à $\operatorname{Im} p$ parallèlement à $\ker p = \ker(u + id_{\mathbb{E}}) = \mathbb{E}_2$. Un vecteur x est dans $\operatorname{Im} p$ ssi $p(x) = x$, soit $u(x) = x$. On a donc bien $\operatorname{Im} p = \ker(u - id_{\mathbb{E}})$.

■

Nous reviendrons sur ce genre d'équation polynomiale dans $\mathcal{L}(\mathbb{E})$ dans le chapitre sur la réduction des endomorphismes. Remarquons que, dans l'étude précédente, on a aussi

$$\mathbb{E}_1 = \mathrm{Im}(u + id_E) \text{ et } \mathbb{E}_2 = \mathrm{Im}(u - id_{\mathbb{E}})$$

et que la décomposition d'un vecteur dans la somme directe $\mathbb{E}_1 \oplus \mathbb{E}_2$ est tout simplement

$$x = \frac{1}{2}(x + u(x)) + \frac{1}{2}(x - u(x))$$

2-2.7 Exercices : "théorèmes" de factorisation

Les exercices qui suivent sont assez instructifs, car ils illustrent bien la manière dont on peut utiliser le recollement d'applications linéaires pour construire des morphismes vérifiant des conditions imposées.

EXERCICE 2-2.17 On se donne trois espaces vectoriels $\mathbb{E}, \mathbb{F}$ et $\mathbb{G}$, et des morphismes $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ et $v \in \mathcal{L}(\mathbb{E}, \mathbb{G})$. On recherche une condition nécessaire et suffisante pour que l'on puisse factoriser v par u à droite, c'est-à-dire pour que l'on puisse trouver un morphisme $w : \mathbb{F} \to \mathbb{G}$ rendant le diagramme

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{u} & \mathbb{F} \\ v\downarrow & \swarrow w & \\ \mathbb{G} & & \end{array}$$

commutatif, donc pour que l'on ait $v = w \circ u$. On admet ici que tout sous-espace vectoriel admet au moins un supplémentaire.

Indication : une condition nécessaire est évidemment $\ker u \subset \ker v$. On montre par analyse-synthèse que cette condition est suffisante. Si w répond à la question, la restriction de w à $\mathrm{Im}\, u$ est imposée : à un vecteur de la forme $u(x)$, w doit **nécessairement** associer le vecteur $v(x)$. Montrer que la condition d'inclusion des noyaux permet de définir sans ambiguïté un morphisme w_1 de $\mathrm{Im}\, u$ dans $\mathrm{Im}\, v$. Pour définir un morphisme w de $\mathbb{F}$ vers $\mathbb{G}$, considérer ensuite une décomposition

$$\mathbb{F} = \mathrm{Im}\, u \oplus \mathbb{S}$$

et recoller w_1 avec ce qui tombe sous la main (et qui soit linéaire quand même! l'application nulle est toujours un bon candidat). On peut montrer que w est unique ssi u est surjectif.

EXERCICE 2-2.18 On se donne trois espaces vectoriels $\mathbb{E}$, $\mathbb{F}$ et $\mathbb{G}$, et des morphismes $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ et $v \in \mathcal{L}(\mathbb{G}, \mathbb{F})$. On recherche une condition nécessaire et suffisante pour que l'on puisse factoriser v par u à gauche, c'est-à-dire pour que l'on puisse trouver un morphisme $w : \mathbb{G} \to \mathbb{E}$ rendant le diagramme

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{u} & \mathbb{F} \\ w\uparrow & \nearrow v & \\ \mathbb{G} & & \end{array}$$

commutatif, soit $v = u \circ w$.

Indication : La condition est ici $\operatorname{Im} v \subset \operatorname{Im} u$. Si l'on réfléchit au problème posé, il s'agit de trouver un antécédent y par u à un vecteur de la forme $v(x)$ pour $x \in \mathbb{G}$, telle que la correspondance $x \mapsto y$ soit linéaire (et c'est cette exigence de **linéarité** qui nécessite réflexion). Si $\mathbb{S}$ est un supplémentaire de $\ker u$ dans $\mathbb{E}$, on sait que u induit un isomorphisme $\psi : \mathbb{S} \to \operatorname{Im} u$. $\psi^{-1} \circ v$ n'est-il pas un candidat sympathique ? Ici encore, on peut montrer que w est unique ssi u est injectif.

Une fois ces résultats acquis, utilisons les pour caractériser les éléments inversibles à droite ou à gauche dans une algèbre $\mathcal{L}(\mathbb{E})$:

EXERCICE 2-2.19 Si u est un endomorphisme d'un espace vectoriel $\mathbb{E}$ montrer l'équivalence des propriétés :

1. u est injectif.
2. u est régulier à gauche (c'est-à-dire pour v et $w \in \mathcal{L}(\mathbb{E})$ $uv = uw \Rightarrow v = w$)
3. u n'est pas diviseur de zéro à gauche ($uv = 0 \Rightarrow v = 0$)
4. u est inversible à gauche ($\exists\, v \in \mathcal{L}(\mathbb{E}) \quad vu = id_{\mathbb{E}}$)

On remarquera que chercher un inverse à gauche revient à essayer de factoriser $id_{\mathbb{E}}$ à droite par u. On pourra montrer de plus que, si ces conditions sont vérifiées et si l'inverse à gauche est unique, alors u est un automorphisme.

EXERCICE 2-2.20 Donner l'énoncé d'un exercice caractérisant les endomorphismes surjectifs d'un espace vectoriel $\mathbb{E}$.

EXERCICE 2-2.21 Résoudre l'exercice précédent.

Fort heureusement, lorsque l'espace $\mathbb{E}$ est **de dimension finie**, les distinctions subtiles entre la droite et la gauche ne sont plus de mise. C'est en partie l'objet de la section qui suit.

2-3 Morphismes et dimension finie

2-3.1 Rang d'un morphisme

DÉFINITION 2-3.1 *On dit qu'un morphisme* $u : \mathbb{E} \to \mathbb{F}$ *est de rang fini si et seulement si* $\operatorname{Im} u$ *est un espace vectoriel de dimension finie. On pose alors, par définition*

$$\operatorname{rang} u = \dim \operatorname{Im} u$$

Par convention, on posera $\operatorname{rang} u = +\infty$ *dans le cas où* u *n'est pas de rang fini.*

Les propriétés suivantes résultent immédiatement de la définition du rang d'un morphisme et d'un système de vecteurs.

- Si $\mathcal{F}$ est un famille génératrice de $\mathbb{E}$, $u(\mathcal{F})$ engendre $\operatorname{Im} u$ et donc

$$\operatorname{rang} u = \operatorname{rang} u(\mathcal{F})$$

  Ceci est en particulier vrai si $\mathcal{F}$ est une base de $\mathbb{E}$.

- Si $\mathbb{E}$ ou $\mathbb{F}$ est de dimension finie, tout morphisme $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ est de rang fini et

$$\operatorname{rang} u \leqslant \dim \mathbb{E} \qquad \operatorname{rang} u \leqslant \dim \mathbb{F}$$

  Si $\mathbb{E}$ est de dimension finie, la première inégalité est une égalité ssi u est injectif. Si $\mathbb{F}$ est de dimension finie, la deuxième inégalité est une égalité ssi u est surjectif.

- Comme conséquence du corollaire **2-1**.23, si l'un des espaces $\mathbb{E}$ ou $\mathbb{F}$ est de dimension finie, un morphisme $u : \mathbb{E} \to \mathbb{F}$ est un isomorphisme ssi

$$\operatorname{rang} u = \dim \mathbb{E} = \dim \mathbb{F}$$

THÉORÈME 2-3.2 Si $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ et $v \in \mathcal{L}(\mathbb{F}, \mathbb{G})$, on a

$$\operatorname{rang}(v \circ u) \leqslant \inf(\operatorname{rang} u\,,\, \operatorname{rang} v)$$

Démonstration : cette majoration n'a évidemment d'intérêt que si l'un au moins des morphismes u ou v est de rang fini. Si u est de rang fini, si $\mathcal{B}$ est une base de $\operatorname{Im} u$, $v(\mathcal{B})$ engendre $\operatorname{Im}(v \circ u)$ (car $\operatorname{Im}(v \circ u) = v(\operatorname{Im} u)$) et donc

$$\operatorname{rang}(v \circ u) \leqslant \operatorname{card}(\mathcal{B}) = \operatorname{rang} u$$

De même, si v est de rang fini, comme $\operatorname{Im}(v \circ u)$ est un sous-espace vectoriel de $\operatorname{Im} v$

$$\operatorname{rang}(v \circ u) \leqslant \operatorname{rang} v \quad \blacksquare$$

COROLLAIRE 2-3.3 Si l'un au moins des facteurs d'un composé de morphismes est de rang fini, ce composé lui-même est de rang fini.

COROLLAIRE 2-3.4 On ne change pas le rang d'un morphisme en le composant (à droite ou à gauche) par un isomorphisme d'espaces vectoriels.

Démonstration : Par exemple, si v est un isomorphisme qu'on peut composer à gauche par u

$$\operatorname{rang}(u \circ v) \leqslant \operatorname{rang} u$$

d'après ce qui précède. Mais comme $u = (u \circ v) \circ v^{-1}$ on a aussi

$$\operatorname{rang}(u) \leqslant \operatorname{rang}(u \circ v) \quad \blacksquare$$

Enfin, le théorème fondamental d'isomorphisme **2-1**.36 peut se ré-écrire (de façon bien compliquée) :

THÉORÈME 2-3.5 Si u est un endomorphisme de $\mathbb{E}$ vers $\mathbb{F}$, u est de rang fini si et seulement si $\ker u$ est de codimension finie dans $\mathbb{E}$ et on a alors

$$\operatorname{rang} u = \operatorname{codim} \ker u$$

Démonstration : Si $\ker u$ est de codimension finie dans $\mathbb{E}$, il possède un supplémentaire de dimension finie $\mathbb{S}$ et on sait que u induit un isomorphisme de $\mathbb{S}$ vers $\operatorname{Im} u$, qui est donc de dimension finie égale à $\dim \mathbb{S} = \operatorname{codim} \ker u$. Réciproquement, si u est de rang fini égal à p, il n'y a rien à dire si u est nul. Sinon, on choisit une base $(u_i)_{1 \leqslant i \leqslant p}$ de $\operatorname{Im} u$ et des vecteurs $(e_i)_{1 \leqslant i \leqslant p}$ de $\mathbb{E}$ tels que $u(e_i) = u_i$. Le système $(e_i)_{1 \leqslant i \leqslant p}$ est libre (car son image par u l'est) et engendre un sous-espace $\mathbb{S}$ de $\mathbb{E}$ qui est un supplémentaire de $\ker u$: en effet, comme l'image d'une base de $\mathbb{S}$ est une base de $\operatorname{Im} u$, u induit un isomorphisme de $\mathbb{S}$ vers $\operatorname{Im} u$ donc $\mathbb{S} \cap \ker u = \{0_E\}$. De plus, si x est un vecteur quelconque de $\mathbb{E}$, $u(x)$ possède un (unique) antécédent dans $\mathbb{S}$ qu'on note y. L'écriture $x = y + (x - y)$ décompose x dans $\mathbb{S} \oplus \ker u$. On a donc $\mathbb{E} = \mathbb{S} \oplus \ker u$ et $\ker u$ est bien de codimension finie. ∎

2-3.2 Théorème du rang

Le théorème fondamental d'isomorphisme entraîne le résultat très important suivant, qui est la réécriture en langage basique du théorème précédent, lorsque l'espace vectoriel de départ est de dimension finie :

THÉORÈME 2-3.6 Si $\mathbb{E}$ est un espace de dimension finie et $u : \mathbb{E} \to \mathbb{F}$ est un morphisme

$$\dim \ker u + \operatorname{rang} u = \dim \mathbb{E}$$

Tout supplémentaire de $\ker u$ dans $\mathbb{E}$ est en effet de dimension égale à la différence $\dim \mathbb{E} - \dim \ker u$.

COROLLAIRE 2-3.7 Si $\mathbb{E}$ et $\mathbb{F}$ sont deux espaces de même dimension finie, un morphisme u de $\mathbb{E}$ dans $\mathbb{F}$ est injectif si et seulement s'il est surjectif.

Démonstration : u est injectif ssi $\dim \ker u = 0$, ce qui équivaut à

$$\operatorname{rang} u = \dim \mathbb{E} = \dim \mathbb{F} < +\infty$$

soit $\dim \operatorname{Im} u = \dim \mathbb{F}$ ou encore $\operatorname{Im} u = \mathbb{F}$, ce qui traduit la surjectivité de u. ∎

Pour prouver que $u : \mathbb{E} \to \mathbb{F}$ est un isomorphisme, après avoir montré que $\dim \mathbb{E} = \dim \mathbb{F} < +\infty$ (et bien évidemment en invoquant la linéarité de u), il suffira donc de prouver l'injectivité (ou la surjectivité) de u. Ceci s'applique évidemment aux endomorphismes d'un espace de dimension finie.

Attention ! Lorsque $\mathbb{E} \neq \mathbb{F}$, il ne suffit pas de dire "on est en dimension finie donc ..." ! Si on ne précise pas qu'espace de départ et d'arrivée ont même dimension finie, on ne prouve rien.

En dimension infinie, un endomorphisme injectif n'est pas surjectif en général. De même la surjectivité n'entraîne pas l'injectivité. Considérer par exemple les morphismes u et $v : \mathbb{R}[\mathbb{X}] \to \mathbb{R}[\mathbb{X}]$ définis par $u(P) = P'$ et $v(P) = XP$.

COROLLAIRE 2-3.8 Si un endomorphisme u d'un espace vectoriel $\mathbb{E}$ de dimension finie est inversible à droite (ou à gauche) dans l'algèbre $\mathcal{L}(\mathbb{E})$, u est un automorphisme et l'inverse à droite (ou à gauche) est égal à u^{-1}

Démonstration : L'inversibilité à droite entraîne évidemment la surjectivité, l'inversibilité à gauche entraînant l'injectivité. Si de plus $u \circ v = id_{\mathbb{E}}$, en composant à gauche par u^{-1}(dont l'existence est assurée par le corollaire **2-3**.7) on obtient immédiatement $v = u^{-1}$. ■

Ici encore, ce corollaire n'est pas valable en dimension infinie. Avec le contre-exemple donné au-dessus dans l'espace $\mathbb{R}[\mathbb{X}]$, si a est un réel quelconque, l'application $\mathbb{R}[\mathbb{X}] \to \mathbb{R}[\mathbb{X}]$ définie par $P \mapsto Q$ avec $Q(x) = \int_a^x P(t)\,dt$ est linéaire et inverse à droite de l'endomorphisme $u : P \mapsto P'$. u n'est cependant pas bijectif, il n'y a de plus pas unicité de l'inverse à droite.

EXERCICE 2-3.9 Si u est un morphisme de $\mathbb{E}$ vers $\mathbb{F}$, espaces vectoriels de dimension finie, démontrer que, pour $\mathbb{E}_1$ et $\mathbb{F}_1$ sous-espaces respectifs de $\mathbb{E}$ et $\mathbb{F}$, on a

$$\begin{aligned} \dim u(\mathbb{E}_1) &= \dim \mathbb{E}_1 - \dim(\mathbb{E}_1 \cap \ker u) \\ \dim u^{-1}(\mathbb{F}_1) &= \dim(\mathbb{F}_1 \cap \operatorname{Im} u) + \dim \ker u \end{aligned}$$

(il suffit d'appliquer le théorème du rang à des restrictions bien choisies de u)

EXERCICE 2-3.10 Si $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$ et $v \in \mathcal{L}(\mathbb{F},\mathbb{G})$ où $\mathbb{E}$, $\mathbb{F}$ et $\mathbb{G}$ sont de dimensions finies, montrer que

$$\operatorname{rang} vu \geqslant \operatorname{rang} u + \operatorname{rang} v - \dim \mathbb{F}$$

2-3.3 Dimension de $\mathcal{L}(\mathbb{E},\mathbb{F})$

Si $\mathbb{E}$ ou $\mathbb{F}$ est réduit au vecteur nul, l'espace $\mathcal{L}(\mathbb{E},\mathbb{F})$ est évidemment réduit à la fonction nulle. On suppose donc dans ce qui suit que ce n'est pas le cas.

THÉORÈME 2-3.11 Si $\mathbb{E}$ et $\mathbb{F}$ sont deux espaces de dimension finie, il en est de même de $\mathcal{L}(\mathbb{E},\mathbb{F})$ et on a

$$\dim \mathcal{L}(\mathbb{E},\mathbb{F}) = \dim \mathbb{E} \times \dim \mathbb{F}$$

Démonstration : Si $\mathcal{B} = (e_j)_{1 \leqslant j \leqslant n}$ est une base de $\mathbb{E}$, le théorème **2-1**.11 dit en substance que l'application

$$\varphi_{\mathcal{B}} : \mathcal{L}(\mathbb{E},\mathbb{F}) \to \mathbb{F}^n \text{ définie par } u \mapsto (u(e_1), \ldots, u(e_n))$$

est un isomorphisme d'espaces vectoriels (la linéarité de cette application est une évidence). Comme la dimension de $\mathbb{F}^n$ est $n \times \dim \mathbb{F}$, le théorème en découle. ■

Le choix d'une base $\mathcal{B}' = (e'_i)_{1\leqslant i\leqslant p}$ de $\mathbb{F}$ permet d'obtenir une base de $\mathbb{F}^n$: il suffit de considérer les np vecteurs $(x_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ définis par

$$x_{ij} = (0_{\mathbb{F}}, \ldots, 0_{\mathbb{F}}, \underset{j^{\text{ème}}\text{place}}{\underset{\downarrow}{e'_i}}, 0_{\mathbb{F}}, \ldots, 0_{\mathbb{F}})$$

En "remontant" cette base par l'isomorphisme $\varphi_{\mathcal{B}}^{-1}$, on obtiendra une base de $\mathcal{L}(\mathbb{E},\mathbb{F})$:

THÉORÈME 2-3.12 Si $\mathcal{B} = (e_j)_{1\leqslant j\leqslant n}$ et $\mathcal{B}' = (e'_i)_{1\leqslant i\leqslant p}$ sont des bases respectives de $\mathbb{E}$ et $\mathbb{F}$, les np morphismes $(u_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ de $\mathbb{E}$ vers $\mathbb{F}$ caractérisés par les images des vecteurs de $\mathcal{B}$:

$$u_{ij}(e_k) = \begin{cases} 0_{\mathbb{F}} & \textbf{si } k \neq j \\ e'_i & \textbf{si } k = j \end{cases}$$

forment une base de $\mathcal{L}(\mathbb{E},\mathbb{F})$.

Cette base dépend du choix des bases de $\mathbb{E}$ et $\mathbb{F}$. Il n'y a pas de base "canonique" dans $\mathcal{L}(\mathbb{E},\mathbb{F})$. On dit cependant souvent que **la base des $(u_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ est la base de $\mathcal{L}(\mathbb{E},\mathbb{F})$ canoniquement associée aux bases $\mathcal{B}$ de $\mathbb{E}$ et $\mathcal{B}'$ de $\mathbb{F}$.** En utilisant le symbole de Kronecker $\delta_k^l = 0$ si $k \neq l$, $= 1$ sinon, on a

$$u_{ij}(e_k) = \delta_k^j e'_i$$

Que signifie pratiquement le théorème **2-3**.11 ? Qu'il faut np scalaires "indépendants" pour repérer un élément de $\mathcal{L}(\mathbb{E},\mathbb{F})$. Si les bases de $\mathbb{E}$ et $\mathbb{F}$ sont choisies comme précédemment, un morphisme $u \in \mathcal{L}(\mathbb{E},\mathbb{F})$ est caractérisé en effet par la donnée des images des vecteurs de $\mathcal{B}$, c'est-à-dire par np scalaires $(\alpha_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ qui donnent les coordonnées de ces images dans $\mathcal{B}'$:

$$\forall j \in \{1,\ldots,n\} \qquad u(e_j) = \sum_{i=1}^{p} \alpha_{ij} e'_i$$

Il est alors facile de vérifier (toujours en raisonnant sur les images des vecteurs de base) que l'on a :

$$u = \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \alpha_{ij} u_{ij}$$

Tout est prêt pour faire le lien entre matrices et morphismes en dimension finie.

2-3.4 Cas particulier de l'espace dual

L'étude de la dualité est l'objet du chapitre suivant. Nous pouvons cependant déjà appliquer le résultat de la section précédente dans le cas où $\mathbb{F} = \mathbb{K}$. La base "naturelle" de $\mathbb{F}$ est alors $\mathcal{B}' = \{1\}$.

DÉFINITION 2-3.13 *Si $\mathbb{E}$ est un $\mathbb{K}$-espace vectoriel, on appelle espace dual de $\mathbb{E}$ l'espace $\mathcal{L}(\mathbb{E}, \mathbb{K})$. On note habituellement cet espace $\mathbb{E}^*$. Un élément de $\mathbb{E}^*$ est appelé* ***forme linéaire*** *sur l'espace $\mathbb{E}$.*

Une forme linéaire est donc une application linéaire qui à un vecteur associe un **scalaire** : l'espace d'arrivée est dans ce cas le plus "simple" possible puisqu'il est de dimension 1.

EXEMPLE 2-3.14 Si α et β sont deux scalaires, l'application de $\mathbb{K}^2$ dans $\mathbb{K}$

$$(x, y) \mapsto \alpha x + \beta y$$

est évidemment une forme linéaire sur $\mathbb{K}^2$.

EXEMPLE 2-3.15 Pour travailler (fort provisoirement) en dimension infinie, deux exemples :

$f \mapsto \int_0^1 f(t)\, dt$ sur l'espace des fonctions réelles continues sur $[0, 1]$.
$P \mapsto P(a)$ sur $\mathbb{K}[X]$, si $a \in \mathbb{K}$ est fixé.

Le théorème **2-3**.12 appliqué dans le cas particulier $\mathbb{F} = \mathbb{K}$ et $\mathcal{B}' = \{1\}$ donne immédiatement :

THÉORÈME 2-3.16 Si $\mathbb{E}$ est un espace de dimension finie n, l'espace dual $\mathbb{E}^*$ est aussi de dimension n. Plus précisément, si $\mathcal{B} = (e_j)_{1 \leqslant j \leqslant n}$ est une base de $\mathbb{E}$, les n formes linéaires $(\varphi_j)_{1 \leqslant j \leqslant n}$ sur $\mathbb{E}$ caractérisées par :

$$\varphi_j(e_k) = \delta_j^k$$

forment une base de $\mathbb{E}^*$.

Si $x = \sum_{i=1}^{n} x_i e_i$ est un vecteur de $\mathbb{E}$ on a :

$$\varphi_j(x) = \sum_{i=1}^{n} x_i \varphi_j(e_i) = x_j$$

Les formes $(\varphi_j)_{1 \leqslant j \leqslant n}$ sont donc **les formes coordonnées dans la base $\mathcal{B}$.** Elles forment une base de $\mathbb{E}^*$ qu'on appelle **base duale** de la base $\mathcal{B}$ et qu'on note parfois $\mathcal{B}^*$. Toute forme linéaire sur l'espace $\mathbb{E}$ s'écrit (de manière unique) comme combinaison linéaire de ces formes coordonnées. Si φ est une forme linéaire quelconque sur $\mathbb{E}$, il existe donc des scalaires (uniques) $(\alpha_j)_{1 \leqslant j \leqslant n}$ tels que

$$\varphi = \sum_{j=1}^{n} \alpha_j \varphi_j$$

Si $x = \sum_{i=1}^{n} x_i e_i$ est dans $\mathbb{E}$, on aura alors

$$\varphi(x) = \sum_{j=1}^{n} \alpha_j x_j$$

Telle est l'expression générale d'une forme linéaire sur un espace de dimension n, **une fois choisie une base $\mathcal{B}$ de $\mathbb{E}$**, choix qui permet d'identifier un vecteur $x \in \mathbb{E}$ avec le n-uplet $(x_1, \ldots, x_n)$ de ses coordonnées dans $\mathcal{B}$. On a l'habitude d'écrire $\varphi(x) = \sum_{j=1}^{n} \alpha_j x_j$ comme le produit

$$\varphi(x) = (\alpha_1, \ldots, \alpha_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

de la **ligne** des coordonnées de φ dans la base $\mathcal{B}^*$ par la **colonne** des coordonnées de x dans la base $\mathcal{B}$.

Bien entendu, φ et x étant donnés, si on change de base $\mathcal{B}$, les x_i changent et bien évidemment aussi les α_i. Comment se transforment-ils ? On le verra dans le chapitre consacré au calcul matriciel.

2-4 Exercices

EXERCICE 2-4.1 Soit $E = \mathbb{R}_n[X]$, A et $B \in E$, $B \neq 0$ et $\Phi : E \to E$ définie par

$$\Phi(P) = \text{Reste de la division euclidienne de } PA \text{ par } B$$

Montrer que $\Phi \in \mathcal{L}(E)$ et déterminer Ker Φ et Im Φ.

EXERCICE 2-4.2 Soit E un $\mathbb{R}$-ev. On dit que E est précomplexe s'il existe sur E une structure de $\mathbb{C}$-ev qui, par restriction du corps des scalaires, induit la structure de $\mathbb{R}$-ev.

1. Montrer que : E précomplexe $\Longleftrightarrow (\exists \phi \in L_{\mathbb{R}}(E)) \quad \phi^2 = -id_E$
2. Montrer que si E est de dimension finie, il est précomplexe ssi sa dimension est paire.

EXERCICE 2-4.3 Soit E un K-ev et $f \in \mathcal{L}(E)$:

1. Montrer que : $((\exists!\, g \in \mathcal{L}(E)) \quad fg = id_E) \Longrightarrow f \in \mathcal{GL}(E)$.
2. Trouver une CNS pour qu'il existe g et $h \in \mathcal{L}(E)$ avec $f = gh$ et $hg = 0$.
3. Montrer que si $\dim E < +\infty$, f peut se factoriser sous la forme $f = gp$ avec $g \in \mathcal{GL}(E)$ et p projecteur.
4. Montrer que $E = \text{Ker} f + \text{Im} f \Longleftrightarrow \text{Im} f = \text{Im} f^2$. Caractériser de même la condition d'indépendance linéaire du noyau et de l'image de f. Qu'obtient-on en dimension finie ?

EXERCICE 2-4.4 Soit E un K-ev et $f, g \in \mathcal{L}(E)$ avec $fg - gf = id_E$.

1. Montrer que pour $P \in K[X]$ on a $P(f)g - gP(f) = P'(f)$.
2. Montrer que cela n'est pas possible si E est de dimension finie.
3. Donner un exemple, en dimension infinie, d'un tel couple d'endomorphismes.

EXERCICE 2-4.5 Soit $E = \mathbf{R}_n[X]$ et $h_0, h_1, \ldots, h_n$ $n+1$ réels deux à deux distincts.

1. Montrer que $\left[(X+h_i)^n\right]_{i=0\ldots n}$ est une base de E.
2. Montrer que l'ensemble des $\varphi \in \mathcal{L}(E)$ commutant aux translations, c'est à dire vérifiant :

$$\forall h \in \mathbf{R} \ \ \forall P \in E \ \ [\varphi(P)](X+h) = \varphi(P(X+h))$$

   est une sous-algèbre de $\mathcal{L}(E)$ de dimension $n+1$.

EXERCICE 2-4.6 Soient $u, v, w \in \mathcal{L}(E)$. Montrer que :

$$\operatorname{Ker}(u) \cap \operatorname{Ker}(v) \subset \operatorname{Ker}(w) \iff \exists\, a, b \in \mathcal{L}(E) \quad w = au + bv$$

Enoncer et démontrer un résultat analogue concernant les images.

EXERCICE 2-4.7 Soit E un espace vectoriel de dimension n et f et g deux endomorphismes de E tels que

$$f + g = id_E \text{ et } \operatorname{rg}(f) + \operatorname{rg}(g) \leqslant n$$

Montrer que f et g sont deux projecteurs. Généralisation : si

$$f_1 + \cdots + f_p = id_E \text{ et } \sum \operatorname{rg}(f_i) \leqslant n$$

alors les f_i sont des projecteurs, vérifient $f_i f_j = 0$ si $i \neq j$ et $E = \bigoplus \operatorname{Im} f_i$. Montrer que si le corps de base est de caractéristique nulle, une somme de projecteurs $p_1 + \cdots + p_l$ est un projecteur ssi

$$\forall i \neq j \ \ p_i p_j = 0$$

(On remarquera que la trace d'un projecteur est égal à son rang).

EXERCICE 2-4.8 Idéaux de $\mathcal{L}(E)$: Soit E un K-ev de dimension finie. On note $\mathcal{P}$ le sous-ensemble de $\mathcal{L}(E)$ formé des projecteurs. Si F est un sev de E, on pose

$$g(F) = \{u \in \mathcal{L}(E) \mid F \subset \operatorname{Ker}(u)\}$$

Si H est un sous-ensemble de $\mathcal{L}(E)$ on note

$$f(H) = \bigcap_{u \in H} \operatorname{Ker}(u)$$

1. Montrer que, pour F sev de E, $g(F)$ est un idéal à gauche de $\mathcal{L}(E)$ et que $f[g(F)] = F$. On désigne par la suite par $\mathcal{I}$ un idéal à gauche de $\mathcal{L}(E)$.
2. Montrer que $\forall u \in \mathcal{L}(E) \ \ \forall p \in \mathcal{P} \ \ \operatorname{Ker}(u) = \operatorname{Ker}(p) \Rightarrow \exists v \in \mathcal{GL}(E) \ \ p = v \circ u$.

3. Si p et q sont deux projecteurs appartenant à $\mathcal{I}$, montrer qu'il existe $r \in \mathcal{P} \cap \mathcal{I}$ tel que
$$\text{Ker}(r) = \text{Ker}(p) \cap \text{Ker}(q)$$

4. Montrer qu'il existe un projecteur $p \in \mathcal{I}$ avec $\text{Ker}(p) = f(\mathcal{I})$.

5. En déduire un caractérisation des idéaux à gauche de $\mathcal{L}(E)$. Déterminer de même les idéaux à droite de $\mathcal{L}(E)$.

EXERCICE 2-4.9 Soient E un Kev de dimension finie et $f \in \mathcal{L}(E)$; pour tout $n \in \mathbb{N}$, on définit $N_n = \ker f^n$, $I_n = \text{Im} f^n$ et $d_n = \dim N_{n+1} - \dim N_n$.
Montrer que la suite $(N_n)_{n \geqslant 0}$ est croissante pour l'inclusion, que $(I_n)_{n \geqslant 0}$, elle, est décroissante et enfin que la suite d'entiers $(d_n)_{n \geqslant 0}$ est décroissante.

EXERCICE 2-4.10 E étant un $\mathbb{R}$ev de dimension finie, montrer que $\mathcal{L}(E)$ admet une base formée de projecteurs.

EXERCICE 2-4.11 Soit E un $\mathbb{R}$-ev et $u \in \mathcal{L}(E)$. Trouver une condition nécessaire et suffisante sur u pour qu'il existe un projecteur p tel que $u = p \circ u - u \circ p$.

Chapitre 3

Dualité

3-1 Espace dual, formes linéaires

3-1.1 Crochet de dualité

Rappelons les définitions données à la fin du chapitre précédent :

DÉFINITION 3-1.1 *Si $\mathbb{E}$ est $\mathbb{K}$-un espace vectoriel, l'espace $\mathbb{E}^* = \mathcal{L}(\mathbb{E}, \mathbb{K})$ est appelé espace dual de l'espace $\mathbb{E}$. Ses éléments sont appelés formes linéaires sur l'espace $\mathbb{E}$.*

En dimension n, si une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est fixée, et si $\varphi \in \mathbb{E}^*$ est donnée, il existe des scalaires $(\alpha_1, \ldots, \alpha_n)$ tels que

$$\forall\, x = \sum_{i=1}^{n} x_i e_i \in \mathbb{E} \qquad \varphi(x) = \sum_{i=1}^{n} \alpha_i x_i$$

On représente alors, **lorsqu'on travaille dans la base** $\mathcal{B}$, la forme linéaire φ par la ligne $(\alpha_1, \ldots, \alpha_n)$.

En particulier, en ramenant $\mathbb{K}^n$ à sa base canonique, on obtient la proposition :

PROPOSITION 3-1.2 Toute forme linéaire sur $\mathbb{K}^n$ est de la forme

$$(x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} \alpha_i x_i = (\alpha_1, \ldots, \alpha_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

où $(\alpha_1, \ldots, \alpha_n)$ **est une ligne de scalaires fixés.**

Une forme linéaire sur $\mathbb{K}^n$ **est donc assimilable à une fonction polynôme homogène de degré 1 de** n **variables.**

DÉFINITION 3-1.3 *Si* $\mathbb{E}$ *est un* $\mathbb{K}$*-ev, l'application de* $\mathbb{E} \times \mathbb{E}^*$ *vers* $\mathbb{K}$ *définie par*

$$(x, \varphi) \mapsto \langle x, \varphi \rangle \underset{\text{déf}}{=} \varphi(x)$$

est appelée crochet de dualité (canonique) sur $\mathbb{E} \times \mathbb{E}^*$.

Cette application est **bilinéaire**, ce qui signifie que les applications partielles (obtenues en fixant un des arguments x ou φ et en faisant varier le second) sont des applications linéaires de $\mathbb{E}^* \to \mathbb{K}$ et de $\mathbb{E} \to \mathbb{K}$. (Les linéarités de ces applications provenant de la définition des opérations d'espace vectoriel dans $\mathbb{E}^*$ pour l'une, et du fait que l'on travaille avec des formes linéaires pour l'autre).

REMARQUE 3-1.4 **En dimension finie**, nous venons de rappeler le théorème **2-3**.16 : toute forme linéaire est combinaison linéaire des formes coordonnées dans une base. C'est la notion de base duale. **En dimension infinie, ce n'est plus le cas.** Si $\mathcal{B} = (e_i)_{i \in I}$ est une base (avec I infini), une forme linéaire φ est caractérisée (théorème **2-1**.11) par la donnée d'une famille **quelconque** de scalaires $(\alpha_i)_{i \in I}$ avec $\alpha_i = \varphi(e_i)$. Si un vecteur x est décomposé dans la base $\mathcal{B}$ sous la forme

$$x = \sum_{i \in I} x_i e_i$$

avec les x_i **presque tous nuls**, on aura bien

$$\varphi(x) = \sum_{i \in I} \alpha_i x_i$$

(presque tous les termes de cette somme sont nuls). Il est bien sûr toujours possible de définir les formes coordonnées $\varphi_i : x \mapsto x_i$. Mais si ψ est une forme combinaison linéaire des formes coordonnées $(\varphi_i)_{i \in I}$ dans la base $\mathcal{B}$, une telle combinaison ne fait intervenir qu'un nombre **fini** de ces formes, c'est-à-dire s'écrit

$$\psi = \sum_{j \in I} \beta_j \varphi_j$$

avec les β_j presque tous nuls. On a alors évidemment $\psi(e_i) = \beta_i$. La forme φ considérée plus haut n'est donc dans $\operatorname{vect}(\varphi_i)_{i \in I}$ que si les coefficients α_i sont presque tous nuls. Autrement dit, le choix de la base $\mathcal{B}$ permet d'identifier l'espace $\mathbb{E}^*$ avec l'espace $\mathbb{K}^I$ des familles quelconques de scalaires indexées par I. Dans cette identification, $\operatorname{vect}(\varphi_i)_{i \in I}$ est le sous-espace $\mathbb{K}^{(I)}$ des familles de scalaires presque tous nuls (*cf.* corollaire **1-3**.13).

Par exemple, dans $\mathbb{K}[X]$, qu'on rapporte à la base canonique $\mathcal{B} = (X^n)_{n \in \mathbb{N}}$, la forme

$$\varphi : P \mapsto P(1)$$

vérifie, pour tout n, $\varphi(X^n) = 1$. Elle n'est donc pas combinaison linéaire de la famille des formes coordonnées $(\varphi_n)_{n\in\mathbb{N}}$. Si un polynôme donné s'écrit

$$P = \sum_{n\in\mathbb{N}} a_n X^n$$

on aura

$$\varphi(P) = \sum_{n\in\mathbb{N}} a_n = \sum_{n\in\mathbb{N}} \varphi_n(P)$$

(il s'agit en fait d'une somme finie). Il n'est pas correct d'écrire $\varphi = \sum_{n\in\mathbb{N}} \varphi_n$, car il s'agit là d'une somme d'une famille infinie de formes non nulles, ce qui n'a pas de sens.

EXERCICE 3-1.5 Dans la remarque précédente, il est implicite que les formes coordonnées dans une base de $\mathbb{E}$ forment toujours une famille libre de $\mathbb{E}^*$. Prouver ce résultat.

3-1.2 Equation d'un hyperplan vectoriel

Dans le plan rapporté à un repère $\left(O, \vec{\imath}, \vec{\jmath}\right)$, donc identifié à $\mathbb{R}^2$, il est bien connu qu'une droite passant par l'origine a une équation de la forme

$$\alpha x + \beta y = 0$$

où $(\alpha, \beta) \neq (0,0)$ est un couple donné de $\mathbb{R}^2$. On sait également qu'une autre équation linéaire $\lambda x + \mu y = 0$ représente la même droite ssi le couple non nul (λ, μ) est proportionnel à (α, β). De même dans $\mathbb{R}^3$, $\alpha x + \beta y + \gamma z = 0$ est l'équation d'un plan passant par l'origine (si le triplet (α, β, γ) est non nul).

Nous avons défini géométriquement un hyperplan d'un espace vectoriel $\mathbb{E}$ (voir définition **2-1**.25) comme étant un sous-espace vectoriel de codimension 1. En particulier, si $\mathbb{E}$ est de dimension finie n, un hyperplan est un sous-espace de dimension $n - 1$. On va voir dans ce paragraphe comment on peut aussi définir "analytiquement" (par une équation linéaire) un hyperplan. On suppose dans ce qui suit que $\mathbb{E}$ est un espace vectoriel non réduit au vecteur nul.

THÉORÈME 3-1.6 Si φ est une forme linéaire non identiquement nulle sur un espace vectoriel $\mathbb{E}$, son noyau $\ker\varphi$ est un hyperplan de $\mathbb{E}$. Réciproquement, si $\mathbb{H}$ est un hyperplan de $\mathbb{E}$, il existe une forme linéaire non nulle de $\mathbb{E}^*$ telle que $\mathbb{H} = \ker\varphi$.

Démonstration : Soit $\varphi \in \mathbb{E}^* - \{0_{\mathbb{E}^*}\}$. $\mathrm{Im}\,\varphi$ est un sous-espace non réduit à $\{0\}$ de $\mathbb{K}$. C'est donc $\mathbb{K}$, et il existe $a \in \mathbb{E}$ avec $\varphi(a) = 1$. Les sous-espaces $\ker\varphi$ et $\mathrm{vect}(a)$ sont alors clairement indépendants. Si $x \in \mathbb{E}$, l'écriture

$$x = \varphi(x)\,a + (x - \varphi(x)\,a)$$

est une décomposition de x dans $\text{vect}(a) \oplus \ker \varphi$, ce qui montre que ces deux sous-espaces sont supplémentaires et que $\ker \varphi$ est bien de codimension 1. C'est donc un hyperplan de $\mathbb{E}$.

Réciproquement, si $\mathbb{H}$ est un hyperplan, il existe un vecteur a non nul avec

$$\mathbb{E} = \text{vect}(a) \oplus \mathbb{H}$$

ce qui signifie que, pour tout $x \in \mathbb{E}$, il existe un scalaire λ_x unique et un vecteur $h_x \in \mathbb{H}$ tels que

$$x = \lambda_x a + h_x$$

L'application $\varphi : \mathbb{E} \to \mathbb{K}$ définie par $x \mapsto \lambda_x$ est clairement une forme linéaire sur $\mathbb{E}$ dont le noyau est exactement $\mathbb{H}$. C'est en fait la forme obtenue en recollant linéairement la forme $\lambda a \mapsto \lambda$ sur $\text{vect}(a)$ avec la forme nulle sur $\mathbb{H}$. ∎

Les vecteurs de l'hyperplan $\mathbb{H} = \ker \varphi$ sont caractérisés par l'équation

$$\varphi(x) = 0$$

Le théorème suivant montre que cette équation (linéaire et scalaire) est "unique à un scalaire multiplicatif non nul près"[1].

THÉORÈME 3-1.7 Si φ et ψ sont deux formes linéaires sur $\mathbb{E}$, il y a équivalence entre les propriétés :

***i)* Il existe un scalaire non nul λ tel que $\psi = \lambda\varphi$**

ii) $\ker \varphi = \ker \psi$

Démonstration : Prouvons *ii)* ⇒ *i)*, l'implication réciproque étant évidente. Si φ est nulle, il en est évidemment de même de ψ et $\lambda = 1$ convient. Sinon $\mathbb{H} = \ker \varphi = \ker \psi$ est un hyperplan de $\mathbb{E}$ et il existe $a \in \mathbb{E} - \mathbb{H}$ avec

$$\mathbf{E} = \text{vect}(a) \oplus \mathbb{H}$$

L'égalité

$$\psi = \frac{\psi(a)}{\varphi(a)}\varphi$$

se vérifie alors en "recollant" ce qui se passe sur $\text{vect}(a)$ et sur $\mathbb{H}$. ∎

[1] Les éléments de $\mathbb{H} = \ker \varphi$ sont caractérisés par l'équation $\varphi(x) = 0$, mais comme $\mathbb{K}$ est un corps, ceci équivaut aussi, par exemple, à $[\varphi(x)]^2 = 0$. L'unicité (à un scalaire près) de l'équation d'un hyperplan s'entend comme unicité de l'équation **linéaire scalaire**. C'est ainsi que nous retrouverons, lorsque nous étudierons les surfaces du second degré dans un espace affine euclidien de dimension 3, les plans comme cas particuliers de surfaces du second degré alors qu'un plan est en général représenté par une équation du premier degré (une équation affine).

REMARQUE 3-1.8 Le théorème précédent réalise donc une bijection entre l'ensemble $\mathcal{H}$ des hyperplans d'un espace vectoriel $\mathbb{E}$ et l'ensemble $\mathcal{D}$ des droites vectorielles de l'espace dual $\mathbb{E}^*$. A un hyperplan $\mathbb{H}$ de $\mathbb{E}$ correspond la droite vectorielle $\mathbb{K}.\varphi$ où φ est une forme linéaire telle que $\ker \varphi = \mathbb{H}$.

Les exercices suivants ont des solutions très simples en dimension finie (essentiellement à l'aide du théorème de la base incomplète). Pour les résoudre en toute généralité, on admettra le théorème d'existence de supplémentaires.

EXERCICE 3-1.9 Montrer qu'un sous-espace vectoriel strict d'un sous-espace $\mathbb{E}$ est toujours inclus dans un hyperplan. En déduire une autre définition des hyperplans d'un espace vectoriel : les hyperplans de $\mathbb{E}$ sont les sous-espaces maximaux (pour la relation d'inclusion) dans l'ensemble des sous-espaces stricts de $\mathbb{E}$.

EXERCICE 3-1.10 Montrer que le vecteur nul est le seul vecteur de $\mathbb{E}$ sur lequel toutes les formes linéaires s'annulent. En particulier, si x et y sont deux vecteurs distincts de $\mathbb{E}$, on peut trouver une forme linéaire φ vérifiant

$$\varphi(x) \neq \varphi(y)$$

(on dit que les formes linéaires séparent les vecteurs de $\mathbb{E}$).

3-2 Etude du rang d'une famille finie de formes

L'idée de base de cette section est la suivante : pour étudier simultanément p formes linéaires $(\psi_i)_{1 \leqslant i \leqslant p}$ sur un espace $\mathbb{E}$, on les "rassemble" en une application (évidemment linéaire) à valeurs dans $\mathbb{K}^p$

$$\Psi : \mathbb{E} \to \mathbb{K}^p \qquad x \mapsto \Psi(x) = (\psi_1(x), \ldots, \psi_p(x))$$

C'est simple et très efficace !

3-2.1 Condition d'indépendance de p formes linéaires

THÉORÈME 3-2.1 La famille $(\psi_i)_{1 \leqslant i \leqslant p}$ est libre dans $\mathbb{E}^*$ si et seulement si l'application

$$\Psi : \mathbb{E} \to \mathbb{K}^p \qquad x \mapsto \Psi(x) = (\psi_1(x), \ldots, \psi_p(x))$$

est surjective.

Démonstration : Si la famille est liée, il existe une forme combinaison linéaire des autres. Sans nuire à la généralité, on peut supposer que c'est ψ_p. (Remarquons que renuméroter les ψ_i revient en fait à composer Ψ avec un automorphisme de $\mathbb{K}^p$ qui permute les coordonnées) :

$$\psi_p = \sum_{i=1}^{p-1} \alpha_i \psi_i$$

Il est alors clair que Ψ n'est pas surjective, puisque par exemple le p-uplet $(0,\dots,0,1)$ ne possède pas d'antécédent par Ψ.

Réciproquement, si Ψ n'est pas surjective, son image est un sous-espace vectoriel strict de $\mathbb{K}^p$. Le théorème de la base incomplète permet alors de prouver l'existence d'un **hyperplan** $\mathbb{H}$ de $\mathbb{K}^p$ contenant $\operatorname{Im}\Psi$. D'après le théorème **3-1**.6, $\mathbb{H}$ peut être défini par une équation linéaire. Il existe des scalaires $(\alpha_1,\dots,\alpha_p)$ non tous nuls tels que

$$(x_1,\dots,x_p) \in \mathbb{H} \Leftrightarrow \sum_{i=1}^{p} \alpha_i x_i = 0$$

Comme $\operatorname{Im}\Psi \subset \mathbb{H}$, on a

$$\forall x \in \mathbb{E} \quad \sum_{i=1}^{p} \alpha_i \psi_i(x) = 0$$

ce qui donne la relation de liaison

$$\sum_{i=1}^{p} \alpha_i \psi_i = 0_{\mathbb{E}^*} \quad \blacksquare$$

Avant d'aller plus loin, voyons la signification de ce théorème lorsqu'on travaille avec $\mathbb{E} = \mathbb{K}^n$. En travaillant dans la base duale de la base canonique, on représente une forme linéaire par une ligne de n scalaires. La donnée des p formes linéaires ψ_i correspond donc à celle de p vecteurs "lignes" de $\mathbb{K}^n$:

$$\text{Pour } i = 1,\dots,p \qquad \psi_i \longleftrightarrow L_i = (\alpha_{i1},\dots,\alpha_{ij},\dots,\alpha_{in})$$

Ces p lignes sont indépendantes si et seulement si, quelle que soit la donnée des scalaires $\beta_1,\dots,\beta_p$, le système d'équations linéaires d'inconnues $(x_1,\dots,x_n)$

$$\begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1j}x_j + \cdots + \alpha_{1n}x_n & = & \beta_1 \\ \vdots & & \vdots \\ \alpha_{i1}x_1 + \cdots + \alpha_{ij}x_j + \cdots + \alpha_{in}x_n & = & \beta_i \\ \vdots & & \vdots \\ \alpha_{p1}x_1 + \cdots + \alpha_{pj}x_j + \cdots + \alpha_{pn}x_n & = & \beta_p \end{cases}$$

possède au moins une solution. Cela correspond bien à l'idée intuitive d'équations indépendantes, chaque ligne, c'est-à-dire chaque forme linéaire, correspondant à une équation.

En toute généralité, l'utilisation du théorème du rang va permettre de déterminer le rang d'une famille finie de formes linéaires. Traitons d'abord un exemple.

3-2.2 Interpolation de Lagrange

Le problème est le suivant : si on se donne p scalaires distincts $x_1, \dots, x_p$ et p scalaires $a_1, \dots, a_p$, trouver un polynôme P de $\mathbb{K}[X]$ tel que

$$\forall i = 1, \dots, p \qquad P(x_i) = a_i$$

C'est un problème d'interpolation : dans le plan $\mathbb{K}^2$, on essaie de faire passer une courbe d'équation polynomiale $y = P(x)$ par les p points M_i de coordonnées (x_i, a_i). Le rapport avec ce qui précède est clair si on remarque que l'application

$$\psi_i : \mathbb{K}[X] \to \mathbb{K} \text{ définie par } Q \mapsto Q(x_i)$$

est une forme linéaire sur $\mathbb{K}[X]$. On est donc en présence d'une équation linéaire et on recherche un antécédent par l'application Ψ (construite comme précédemment) pour le p-uplet $(a_1, \dots, a_p)$.

Si le problème possède une solution, celle-ci n'est pas unique, puisque le noyau de l'application Ψ est l'ensemble des polynômes ayant les x_i comme racines, donc divisibles par le polynôme $\pi(X) = \prod_{i=1}^{p} (X - x_i)$. Deux polynômes solutions diffèrent d'un multiple de π, et par division euclidienne, on peut rechercher les solutions de degré inférieur ou égal à $p - 1$. (Une autre raison, plus intuitive mais beaucoup moins précise qui pourrait amener à travailler dans $\mathbb{K}_{p-1}[X]$ serait qu'on a à résoudre un système de p équations scalaires, et qu'il est naturel de travailler dans $\mathbb{K}_{p-1}[X]$, qui est de dimension p : en travaillant dans cet espace, on cherche à résoudre un problème à p inconnues).

On considère donc la restriction Ψ_p de Ψ à $\mathbb{K}_{p-1}[X]$:

$$\Psi_p : \mathbb{K}_{p-1}[X] \to \mathbb{K}^p \text{ définie par } Q \mapsto (Q(x_1), \dots, Q(x_p))$$

Cette application est linéaire, injective (déterminer son noyau) et les espaces de départ et d'arrivée sont de même dimension finie. Il en découle que Ψ_p est un **isomorphisme** d'espaces vectoriels et on en déduit

THÉORÈME 3-2.2 Si $x_1, \dots, x_p$ sont p éléments distincts de $\mathbb{K}$ et $a_1, \dots, a_p$ sont quelconques dans $\mathbb{K}$, il existe un unique polynôme P de $\mathbb{K}_{p-1}[X]$ tel que

$$\forall i \in \{1, \dots, p\} \quad P(x_i) = a_i$$

COROLLAIRE 3-2.3 Sous les hypothèses précédentes, les p formes linéaires sur $\mathbb{K}_{p-1}[X]$

$$Q \mapsto Q(x_i)$$

forment une base de l'espace dual $(\mathbb{K}_{p-1}[X])^*$

COROLLAIRE 3-2.4 Sous les hypothèses précédentes, les p formes linéaires sur $\mathbb{K}[X]$

$$Q \mapsto Q(x_i)$$

sont indépendantes.

Comment trouver explicitement le polynôme P donné par le théorème précédent ? On pourrait utiliser la méthode "des coefficients indéterminés", c'est-à-dire chercher *a priori* P sous la forme

$$P = \sum_{i=0}^{p-1} \alpha_i X^i$$

et essayer de résoudre le système d'équations correspondant au problème. Il n'est pas impossible de s'en tirer comme cela, mais au prix de calculs complexes. Ce n'est pas la bonne méthode. Pourquoi ? Parce que, ce faisant, on essaie d'exprimer le polynôme P dans la base canonique de $\mathbb{K}_{p-1}[X]$ alors que **cette base n'est pas adaptée** au problème traité !

Pour trouver l'antécédent $P = \Psi_p^{-1}(a_1, \ldots, a_p)$, il est plus simple de décomposer

$$(a_1, \ldots, a_p) = \sum_{i=1}^{p} a_i e_i$$

dans la base canonique de $\mathbb{K}^p$ et de "remonter" cette base par l'isomorphisme Ψ_p^{-1} en une base de $\mathbb{K}_{p-1}[X]$ soit $(L_i)_{1\leqslant i\leqslant p}$ avec $L_i = \Psi_p^{-1}(e_i)$. On aura alors, par linéarité

$$P = \sum_{i=1}^{p} a_i L_i = \sum_{i=1}^{p} P(x_i) L_i$$

L'expression de L_i est fort simple : c'est le polynôme (unique) de $\mathbb{K}_{p-1}[X]$ admettant pour racines les $(x_j)_{j\neq i}$ et vérifiant $L_i(x_i) = 1$. Il est proportionnel à $\prod_{j\neq i}(X - x_j)$ et vaut donc

$$L_i = \prod_{j\neq i} \frac{(X - x_j)}{(x_i - x_j)}$$

COROLLAIRE 3-2.5 Le polynôme P dont l'existence et l'unicité sont assurés par le théorème 3-2.2 est donné par

$$P = \sum_{i=1}^{p} \left(a_i \prod_{j\neq i} \frac{(X - x_j)}{(x_i - x_j)} \right) = \sum_{i=1}^{p} \left(P(x_i) \prod_{j\neq i} \frac{(X - x_j)}{(x_i - x_j)} \right)$$

DÉFINITION 3-2.6 *Les polynômes L_i explicités plus haut sont appelés polynômes interpolateurs de Lagrange pour les points $(x_i)_{1\leqslant i\leqslant p}$. Ils forment donc une base de $\mathbb{K}_{p-1}[X]$.*

EXERCICE 3-2.7 Résoudre le problème d'interpolation de Hermite : trouver les polynômes P vérifiant

$$\forall\, i \in \{1, \ldots, p\} \qquad P(x_i) = a_i \text{ et } P'(x_i) = b_i$$

où les a_i et b_i sont $2p$ scalaires arbitraires.

3-2.3 Rang d'une famille finie de formes linéaires

Les notations sont celles de la section **3-2.1**. Lorsque les formes $(\psi_i)_{1\leqslant i\leqslant p}$ sont indépendantes, l'application $\Psi : \mathbb{E} \to \mathbb{K}^p$ est surjective, donc de rang p. Nous généralisons ici ce résultat. Le théorème qui suit aura une traduction matricielle remarquable : le rang du système de vecteurs colonnes d'une matrice est égal au rang du système des vecteurs lignes, ce qu'on écrira aussi

$$\forall A \in \mathcal{M}_{p,n}(\mathbb{K}) \qquad \operatorname{rang} A = \operatorname{rang} {}^t A$$

(voir la démonstration du théorème **4-1**.16).

THÉORÈME 3-2.8 Le rang de la famille $(\psi_i)_{1\leqslant i\leqslant p}$ dans $\mathbb{E}^*$ est égal au rang de l'application

$$\Psi : \mathbb{E} \to \mathbb{K}^p \qquad x \mapsto \Psi(x) = (\psi_1(x), \ldots, \psi_p(x))$$

Démonstration : Appelons r le rang de la famille $(\psi_i)_{1\leqslant i\leqslant p}$. Si $r = 0$, l'application Ψ est identiquement nulle et est aussi de rang 0. Sinon, nous déterminons le rang de Ψ à l'aide du théorème **2-3**.5. On extrait de la famille $(\psi_i)_{1\leqslant i\leqslant p}$ une famille principale, qu'on suppose, sans nuire à la généralité, égale à $(\psi_i)_{1\leqslant i\leqslant r}$. On a alors

$$\ker \Psi = \bigcap_{i=1}^{p} \ker \psi_i = \bigcap_{i=1}^{r} \ker \psi_i$$

puisqu'en effet, chaque ψ_j pour $j > r$ étant combinaison linéaire de $(\psi_i)_{1\leqslant i\leqslant r}$ est identiquement nulle sur $\bigcap_{i=1}^{r} \ker \psi_i$. On a donc l'égalité $\ker \Psi = \ker \Psi'$ avec

$$\Psi' : \mathbb{E} \to \mathbb{K}^r \qquad x \mapsto (\psi_1(x), \ldots, \psi_r(x))$$

Mais on sait (section **3-2.1**) que cette application est surjective, donc de rang r. Son noyau, égal à $\ker \Psi$, est donc (théorème **2-3**.5) de codimension r, ce qui montre (c'est toujours le même théorème) que Ψ est de rang r. ∎

3-2.4 Intersection d'une famille finie d'hyperplans

Par abus de langage, on dira que des hyperplans $(\mathbb{H}_i)_{1\leqslant i\leqslant r}$ sont indépendants ssi ces hyperplans s'écrivent

$$\mathbb{H}_i = \ker \psi_i$$

où les formes linéaires $(\psi_i)_{1\leqslant i\leqslant r}$ sont indépendantes. Le théorème du paragraphe précédent a alors la traduction géométrique suivante :

THÉORÈME 3-2.9 L'intersection d'une famille de r hyperplans indépendants est un sous-espace vectoriel de codimension r. Plus généralement, l'intersection d'une famille de p hyperplans $\mathbb{H}_i = \ker \psi_i$ est un sous-espace vectoriel de codimension r, où r est le rang de la famille de formes $(\psi_i)_{1\leqslant i\leqslant p}$.

COROLLAIRE 3-2.10 Dans un espace vectoriel de dimension finie n, une intersection de r hyperplans indépendants est un sous-espace vectoriel de dimension $n - r$.

L'exercice suivant est une généralisation de la remarque **3-1**.8 relative à une identification entre les droites vectorielles de $\mathbb{E}^*$ et les hyperplans de $\mathbb{E}$. On va ici identifier les sous-espaces de dimension finie du dual avec les sous-espaces de codimension finie de l'espace :

EXERCICE 3-2.11 On considère un espace vectoriel $\mathbb{E}$ et $\mathbb{E}^*$ son dual.

- Soit $\mathbb{F}$ un sous-espace **de dimension finie** $p > 0$ de $\mathbb{E}^*$ et $(\psi_1, \ldots, \psi_p)$ une base de $\mathbb{F}$. On note

$$^\circ\mathbb{F} = \{x \in \mathbb{E} \mid \forall \varphi \in \mathbb{F} \quad \varphi(x) = 0\}$$

  Montrer que $^\circ\mathbb{F}$ est un sous-espace de codimension p de $\mathbb{E}$, et donc que

$$\operatorname{codim} {}^\circ\mathbb{F} = \dim \mathbb{F}$$

- Soit $\mathbb{G}$ un sous-espace de **codimension finie** $p > 0$ de $\mathbb{E}$. Il existe donc une famille libre de p vecteurs de $\mathbb{E}$ tels que

$$\mathbb{E} = \mathbb{G} \oplus \operatorname{vect}(\epsilon_1, \ldots, \epsilon_p)$$

  On note

$$\mathbb{G}^\circ = \{\psi \in \mathbb{E}^* \mid \forall x \in \mathbb{G} \quad \psi(x) = 0\}$$

  Montrer que $\mathbb{G}^\circ$ est un sous-espace vectoriel de $\mathbb{E}^*$ de dimension p et donc que

$$\dim \mathbb{G}^\circ = \operatorname{codim} \mathbb{G}$$

  (on remarquera qu'une forme linéaire générale sur $\mathbb{E}$ s'obtient en recollant une forme sur $\mathbb{G}$ et une forme sur un supplémentaire).

- On note $\mathcal{C}_p$ l'ensemble des sous-espaces de $\mathbb{E}$ de codimension p et $\mathcal{D}_p$ l'ensemble des sous-espaces de $\mathbb{E}^*$ de dimension p. Montrer que l'application

$$\mathcal{D}_p \to \mathcal{C}_p \text{ définie par } \mathbb{F} \mapsto {}^\circ\mathbb{F}$$

  est bijective et que $\mathbb{G} \mapsto \mathbb{G}^\circ$ est la bijection réciproque.

  (Une partie de ce qui se cache derrière cet exercice est le fait qu'un sous-espace $\mathbb{G}$ de $\mathbb{E}$ de codimension finie p est entièrement caractérisé par la donnée d'une base (finie) de $\mathbb{G}^\circ$, soit $(\psi_i)_{1 \leqslant i \leqslant p}$, c'est-à-dire que

$$x \in \mathbb{G} \Leftrightarrow \begin{cases} \psi_1(x) = 0 \\ \vdots \\ \psi_p(x) = 0 \end{cases}$$

  Les éléments de $\mathbb{G}$ sont caractérisés par un système de p équations linéaires indépendantes. Nous reviendrons ultérieurement sur ce résultat, en nous plaçant dans le cadre plus confortable d'un espace vectoriel $\mathbb{E}$ de dimension finie).

- Avec cette nouvelle vision de la codimension finie, essayer de reprendre l'exercice **2-1**.29.

3-2.5 Faisceau (linéaires) d'hyperplans

Nous avons étudié, dans la section précédente, l'intersection d'une famille finie d'hyperplans. Il est simple de caractériser les hyperplans qui contiennent alors cette intersection :

THÉORÈME 3-2.12 Si $(\psi_i)_{1\leqslant i\leqslant p}$ est une famille finie de formes linéaires sur un espace vectoriel $\mathbb{E}$ et si $\varphi \in \mathbb{E}^*$, on a :

$$\varphi \in \operatorname{vect}(\psi_i)_{1\leqslant i\leqslant p} \Longleftrightarrow \bigcap_{i=1}^{p} \ker \psi_i \subset \ker \varphi$$

ce qui revient à dire que φ est nulle là où toutes les ψ_i s'annulent.

Démonstration : L'implication directe est évidente. On suppose à présent que l'intersection des noyaux des ψ_i est incluse dans le noyau de φ. En posant $\psi_{p+1} = \varphi$, on a alors

$$\bigcap_{i=1}^{p} \ker \psi_i = \bigcap_{i=1}^{p+1} \ker \psi_i$$

Le théorème **3-2**.9 montre alors que

$$\operatorname{rang}(\psi_1, \dots, \psi_p) = \operatorname{rang}(\psi_1, \dots, \psi_p, \varphi)$$

ce qui montre que $\varphi \in \operatorname{vect}(\psi_1, \dots, \psi_p)$. ■

En laissant de côté le cas où l'on travaille avec des formes nulles, le théorème précédent s'énonce de manière plus géométrique :

COROLLAIRE 3-2.13 Si $(\mathbb{H}_i)_{1\leqslant i\leqslant p}$ est une famille finie d'hyperplans d'un espace vectoriel $\mathbb{E}$, un hyperplan $\mathbb{H}$ passe par l'intersection des $\mathbb{H}_i$ si et seulement si une équation de $\mathbb{H}$ est combinaison linéaire des équations des $\mathbb{H}_i$.

EXEMPLE 3-2.14 Recherche de l'équation générale d'un plan passant par une droite donnée. Faisceau de plan d'arête $\mathcal{D}$.

Pour illustrer le résultat précédent, plaçons nous dans un espace affine réel de dimension 3 rapporté à un repère $\mathcal{R} = \left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$. Deux plans affines non parallèles $\mathcal{P}_1$ et $\mathcal{P}_2$ ont pour intersection une droite $\mathcal{D}$. En travaillant dans le repère $\mathcal{R}$, on se donne des équations des $\mathcal{P}_i$:

$$(\mathcal{S}) : \begin{cases} (\mathcal{P}_1) : & a_1x + b_1y + c_1z + d_1 = 0 \\ (\mathcal{P}_2) : & a_2x + b_2y + c_2z + d_2 = 0 \end{cases}$$

En considérant les formes linéaires φ_1 et φ_2 sur l'espace vectoriel sous-jacent-identifié avec $\mathbb{R}^3$ par le choix de la base $\left(\overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$:

$$\varphi_i(x, y, z) = \varphi_i\left(x\overrightarrow{I} + y\overrightarrow{J} + z\overrightarrow{K}\right) = a_ix + b_iy + c_iz$$

le non parallélisme des plans se traduit par l'indépendance des formes linéaires φ_1 et φ_2 (donc des lignes (a_1, b_1, c_1) et (a_2, b_2, c_2)). Remarquons en passant que le théorème **3-2**.1 assure l'existence de solutions au système $(\mathcal{S})$, et que l'ensemble des solutions de $(\mathcal{S})$ est alors un sous-espace affine de $\mathbb{R}^3$ de direction $\ker\varphi_1 \cap \ker\varphi_2$ (voir la section **2-1.7**) qui représente bien la droite $\mathcal{D}$ dans le repère $\mathcal{R}$.

Si M est le point générique de l'espace, de coordonnées (x, y, z) dans le repère $\mathcal{R}$, un système d'équations définissant $\mathcal{D}$ s'écrit

$$M \in \mathcal{D} \Longleftrightarrow \begin{cases} (\mathcal{P}_1) : & \varphi_1\left(\overrightarrow{OM}\right) + d_1 = 0 \\ (\mathcal{P}_2) : & \varphi_2\left(\overrightarrow{OM}\right) + d_2 = 0 \end{cases}$$

Si M_0 est un point quelconque de $\mathcal{D}$, on a donc $d_i = -\varphi_i\left(\overrightarrow{OM_0}\right)$ et les équations de $\mathcal{D}$ s'écrivent aussi :

$$M(x, y, z) \in \mathcal{D} \Longleftrightarrow \begin{cases} (\mathcal{P}_1) : & a_1x + b_1y + c_1z + d_1 = \varphi_1\left(\overrightarrow{M_0M}\right) = 0 \\ (\mathcal{P}_2) : & a_2x + b_2y + c_2z + d_2 = \varphi_2\left(\overrightarrow{M_0M}\right) = 0 \end{cases}$$

soit

$$\overrightarrow{M_0M} \in \ker\varphi_1 \cap \ker\varphi_2$$

Si $\mathcal{P}$ est un plan quelconque dont une équation, dans le repère $\mathcal{R}$ est

$$(\mathcal{P}) : \alpha x + \beta y + \gamma z + \delta = 0$$

on cherche une condition nécessaire et suffisante portant sur cette équation pour que $\mathcal{P}$ contienne $\mathcal{D}$. Si φ est la forme linéaire définie par

$$\varphi(x, y, z) = \varphi\left(x\overrightarrow{I} + y\overrightarrow{J} + z\overrightarrow{K}\right) = \alpha x + \beta y + \gamma z$$

et si le plan passe par M_0, son équation peut s'écrire de même

$$(\mathcal{P}) : \alpha x + \beta y + \gamma z + \delta = \varphi\left(\overrightarrow{M_0M}\right) = 0$$

Si de plus $\mathcal{P}$ contient $\mathcal{D}$, sa direction contient la direction de la droite $\mathcal{D}$ soit

$$\ker\varphi_1 \cap \ker\varphi_2 \subset \ker\varphi$$

Le théorème précédent assure l'existence de deux scalaires λ_1 et λ_2 (non tous nuls) tels que

$$\varphi = \lambda_1\varphi_1 + \lambda_2\varphi_2$$

ce qui traduit exactement le fait qu'une équation de $\mathcal{P}$ est combinaison linéaire des équations de $\mathcal{P}_1$ et $\mathcal{P}_2$:

$$\exists\ (\lambda_1, \lambda_2) \in \mathbb{R}^2 - \{(0,0)\} \quad \begin{cases} \alpha = \lambda_1 a_1 + \lambda_2 a_2 \\ \beta = \lambda_1 b_1 + \lambda_2 b_2 \\ \gamma = \lambda_1 c_1 + \lambda_2 c_2 \\ \delta = \lambda_1 d_1 + \lambda_2 d_2 \end{cases}$$

Cette condition est aussi évidemment suffisante pour que $\mathcal{P} \supset \mathcal{P}_1 \cap \mathcal{P}_2$. Résumons cette étude :

DÉFINITION 3-2.15 (et théorème) *Si $\mathcal{D}$ est une droite d'un espace affine de dimension 3, on appelle* ***faisceau de plans d'arête*** *$\mathcal{D}$ l'ensemble des plans contenant $\mathcal{D}$. Si $\mathcal{P}_1$ et $\mathcal{P}_2$ sont deux plans distincts de ce faisceau, on dit qu'ils forment une base de ce faisceau. L'équation générale d'un plan du faisceau est alors combinaison linéaire des équations de $\mathcal{P}_1$ et $\mathcal{P}_2$.*

On aurait des résultats analogues en travaillant avec l'ensemble des plans passant par un point fixe. Il faudrait alors travailler avec les combinaisons linéaires des équations de 3 plans particuliers. De même, en dimension 2, deux droites concourantes définissent le faisceau des droites passant par leur intersection.

EXERCICE 3-2.16 Dans un espace affine de dimension 3, si deux plans non parallèles d'équations

$$\begin{aligned}(\mathcal{P}_1):\quad \Phi_1(M) &= 0\\ (\mathcal{P}_2):\quad \Phi_2(M) &= 0\end{aligned}$$

ont pour intersection une droite $\mathcal{D}$ et si M_0 est un point de l'espace pris hors de $\mathcal{D}$, quelle est l'équation du plan défini par $\mathcal{D}$ et M_0 ?

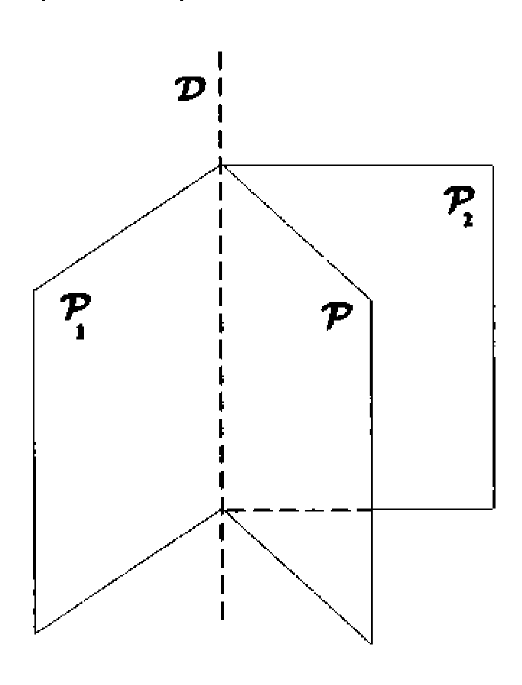


Figure **3.1** – *Faisceau de plans d'arête $\mathcal{D}$*

EXERCICE 3-2.17 Dans un espace affine euclidien de dimension 3 rapporté à **un repère orthonormé**, si deux plans non parallèles ont des équations **normales** (voir le chapitre **22** de rappels de géométrie).

$$\begin{aligned}(\mathcal{P}_1):\quad \Phi_1(M) &= 0\\ (\mathcal{P}_2):\quad \Phi_2(M) &= 0\end{aligned}$$

les **plans bissecteurs** ont pour équations :

$$\Phi_1(M) - \Phi_2(M) = 0 \quad \text{et} \quad \Phi_1(M) + \Phi_2(M) = 0$$

(penser à l'expression de la distance d'un point à un plan).

3-2.6 Bases duales

Attention ! Qui dit base duale dit espace vectoriel **de dimension finie** (voir remarque **3-1**.4). Rappelons la définition de la base duale donnée à la fin du chapitre précédent :

DÉFINITION 3-2.18 *Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est une base d'un espace $\mathbb{E}$ de dimension n, les formes coordonnées dans $\mathcal{B}$ forment une base $\mathcal{B}^*$ de l'espace dual, qu'on appelle base duale de la base $\mathcal{B}$.*

Le théorème qui suit montre qu'en fait toute base de $\mathbb{E}^*$ est de ce type :

THÉORÈME 3-2.19 Si $\mathbb{E}$ est un espace de dimension finie n et $\mathcal{D}$ est une base de $\mathbb{E}^*$, il existe une unique base $\mathcal{B}$ de $\mathbb{E}$ dont $\mathcal{D}$ soit base duale.

Démonstration : Soit $\mathcal{D} = (\psi_i)_{1 \leqslant i \leqslant n}$ une base de $\mathbb{E}^*$. L'application Ψ associée de $\mathbb{E}$ vers $\mathbb{K}^n$ est de rang n, c'est donc un isomorphisme d'espaces vectoriels. L'image réciproque par Ψ de la base canonique de $\mathbb{K}^n$ est donc une base de $\mathbb{E}$, soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, et la condition $\mathcal{B} = \Psi^{-1}(\mathcal{D})$ traduit exactement les conditions

$$\psi_i(e_j) = \delta_i^j$$

ce qui donne donc $\mathcal{D} = \mathcal{B}^*$. L'unicité de $\mathcal{B}$ est conséquence de la traduction du problème avec $\mathcal{B} = \Psi^{-1}(\mathcal{D})$. ■

On dit parfois que $\mathcal{B}$ est la **base "préduale"** de $\mathcal{D}$

Attention ! Le théorème précédent montre que la correspondance $\mathcal{B} \mapsto \mathcal{B}^*$ réalise une bijection entre les bases de $\mathbb{E}$ et celles de $\mathbb{E}^*$. Il ne faudrait pas croire qu'elle permet de réaliser une correspondance entre vecteurs de $\mathbb{E}$ et formes de $\mathbb{E}^*$: si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est une base, la première forme coordonnée dépend **de tous les vecteurs de la base,** et pas seulement de e_1 ! Par exemple, si on ramène l'espace $\mathbb{K}^2$ à sa base canonique (e_1, e_2), la première forme coordonnée est

$$\mathbb{K}^2 \to \mathbb{K} \qquad (x_1, x_2) \mapsto x_1$$

Si on travaille dans la base $(e_1, e_1 + e_2)$, la première forme coordonnée est

$$\mathbb{K}^2 \to \mathbb{K} \qquad (x_1, x_2) \mapsto x_1 - x_2$$

et pourtant, dans les deux cas, le premier vecteur de base est e_1.

EXEMPLE 3-2.20 L'étude du paragraphe **3-2.2** montre que, si $x_1, \ldots, x_n$ sont n scalaires distincts, les formes linéaires $P \mapsto P(x_i)$ forment une base de l'espace dual de $\mathbb{K}_{n-1}[X]$, qui est base duale de la base des polynômes interpolateurs de Lagrange :

$$\mathcal{B} = (L_i)_{1 \leqslant i \leqslant n} \quad \text{avec} \quad L_i = \prod_{j \neq i} \frac{(X - x_j)}{(x_i - x_j)}$$

EXERCICE 3-2.21 Montrer que les formes linéaires

$$P \mapsto P^{(i)}(\alpha) \text{ pour } 0 \leqslant i \leqslant n$$

forment une base du dual de $\mathbb{R}_n[X]$, et trouver la base dont elle est duale.

3-3 Dimension finie : Rang d'une famille et équations d'un s.e.v

3-3.1 Rang d'un système fini de vecteurs et dualité

Nous supposerons ici que l'espace vectoriel $\mathbb{E}$ est de dimension finie, ce qui simplifiera les démonstrations avec notamment la possibilité d'utiliser la notion de base duale (et le théorème du rang). Les résultats de cette section sont cependant valables en dimension infinie, pour peu que l'on admette que le vecteur nul de $\mathbb{E}$ est le seul sur lequel s'annulent **toutes les formes linéaires** (voir l'exercice **3-1.10**) ce qui revient en fait à supposer que toute droite vectorielle possède au moins un (hyperplan) supplémentaire. C'est un bon exercice de reconstruire les démonstrations dans le cas général (sans lire la remarque qui termine cette section !).

L'idée est celle développée en **3-2.1**, en permutant les rôles joués par les formes linéaires et les vecteurs.

THÉORÈME 3-3.1 Si $(e_i)_{1\leqslant i\leqslant p}$ est une famille finie de vecteurs d'un espace $\mathbb{E}$ (de dimension finie n) l'application

$$\Sigma : \mathbb{E}^* \to \mathbb{K}^p \text{ définie par } \varphi \mapsto (\varphi(e_1), \ldots, \varphi(e_p))$$

est surjective si et seulement si la famille $(e_i)_{1\leqslant i\leqslant p}$ est libre.

Démonstration : Cette application est évidemment linéaire. Si la famille $(e_i)_{1\leqslant i\leqslant p}$ est liée, il existe une relation de liaison non triviale

$$\sum_{i=1}^{p} \alpha_i e_i = 0_{\mathbb{E}}$$

Il est alors immédiat de vérifier que l'image de Σ est incluse dans l'hyperplan $\mathbb{H}$ de $\mathbb{K}^p$ d'équation

$$(\mathbb{H}) : \sum_{i=1}^{p} \alpha_i x_i = 0$$

et Σ n'est pas surjective.

Si la famille $(e_i)_{1\leqslant i\leqslant p}$ est libre, on peut la compléter en une base de $\mathbb{E}$ soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$. On note $\mathcal{B}^* = (\varphi_i)_{1\leqslant i\leqslant n}$ la base duale. L'image par Σ de la famille $(\varphi_i)_{1\leqslant i\leqslant p}$ est la base canonique de $\mathbb{K}^p$, ce qui prouve la surjectivité de Σ. ■

COROLLAIRE 3-3.2 Sous les hypothèse précédentes,

$$\operatorname{rang}(e_i)_{1\leqslant i\leqslant p} = \operatorname{rang}\Sigma$$

Démonstration : Rien à dire si ce rang est nul. Sinon, on extrait une famille principale de la famille, qu'on suppose sans nuire à la généralité, égale à $(e_i)_{1\leqslant i\leqslant r}$. Il est alors clair que

$$\ker \Sigma = \{\varphi \in \mathbb{E}^* \mid \forall i \in 1, \dots, p \quad \varphi(e_i) = 0\}$$
$$= \{\varphi \in \mathbb{E}^* \mid \forall i \in 1, \dots, r \quad \varphi(e_i) = 0\}$$

C'est donc le noyau d'un morphisme surjectif

$$\mathbb{E}^* \to \mathbb{K}^r$$

donc de dimension $n - r$. Le théorème du rang montre alors que l'application Σ est effectivement de rang r. ■

COROLLAIRE 3-3.3 Si $(e_i)_{1\leqslant i\leqslant p}$ est une famille finie de vecteurs d'un espace $\mathbb{E}$ (de dimension finie n) un vecteur $x \in \mathbb{E}$ est dans le sous-espace $\mathrm{vect}\,(e_i)_{1\leqslant i\leqslant p}$ si et seulement si toutes les formes linéaires s'annulant sur tous les e_i s'annulent aussi sur x.

Démonstration : Cette condition est évidemment nécessaire. Montrons qu'elle est suffisante : dire que toutes les formes linéaires s'annulant sur tous les e_i s'annulent aussi sur x revient à dire que le noyau de l'application Σ associée à $(e_i)_{1\leqslant i\leqslant p}$ est égal au noyau de l'application Σ' associée à $(e_1, \dots, e_p, x)$. D'après le corollaire précédent, cela donne

$$\mathrm{rang}\,(e_1, \dots, e_p) = \mathrm{rang}\,(e_1, \dots, e_p, x)$$

ce qui traduit bien

$$x \in \mathrm{vect}\,(e_1, \dots, e_p) \quad ■$$

REMARQUE 3-3.4 Nous avons refait les démonstrations de cette section dans un souci pédagogique. En fait, il n'y a pratiquement rien à démontrer, si on accepte de monter d'un étage dans l'édifice :

Si $\mathbb{E}$ est un espace vectoriel, on a défini son espace dual $\mathbb{E}^*$. C'est certes un espace dont les éléments sont des objets "compliqués" (des formes linéaires). Considérons-le cependant comme un espace "ordinaire" et pour ce faire, donnons lui un nom "ordinaire" $\mathbb{F} = \mathbb{E}^*$. On peut alors considérer son dual $\mathbb{F}^*$ (qu'on appelle bidual de $\mathbb{E}$, et qu'on note en général $\mathbb{E}^{**}$). Dans ce dual, considérons des applications d'un type particulier[2] : on se donne un vecteur x de $\mathbb{E}$ et on considère l'application

$$\widetilde{x} : \mathbb{F} \to \mathbb{K} \text{ définie par } \varphi \mapsto \varphi(x)$$

[2]Si l'espace $\mathbb{E}$ est de dimension finie n, il en est de même du bidual $\mathbb{F}^* = \mathbb{E}^{**}$. L'application injective $x \mapsto \widetilde{x}$ est donc un isomorphisme entre $\mathbb{E}$ et $\mathbb{E}^{**}$, isomorphisme qu'on dit "canonique" car il est indépendant du choix de toute base. Si on choisit cependant une base $\mathcal{B}$ de $\mathbb{E}$ et qu'on utilise la base duale $\mathcal{B}^*$ pour décomposer les formes linéaires, il est d'usage de représenter les vecteurs de $\mathbb{E}$ par des colonnes et les formes par des lignes. "Identifier" par $x \leftrightarrow \widetilde{x}$ l'espace et son bidual revient à écrire les coordonnées de x dans $\mathcal{B}$ en lignes et les coordonnées des formes linéaires sur $\mathbb{E}$ en colonnes : les vecteurs deviennent des formes et les formes des vecteurs ! Remarquons que nous venons de redémontrer le théorème **3-2**.19. Lorsque $\mathbb{E}$ est de dimension infinie, l'application $x \mapsto \widetilde{x}$ n'est jamais surjective (exercice en supposant que $\mathbb{E}$ possède une base). Le bidual est "beaucoup plus gros" que l'espace.

C'est évidemment un élément de $\mathbb{F}^*$ et il est facile de voir que la correspondance

$$\mathbb{E} \to \mathbb{F}^* \quad x \mapsto \widetilde{x}$$

est linéaire. Le fait d'admettre que pour un vecteur non nul donné il existe une forme prenant une valeur non nulle sur ce vecteur (résultat clair en dimension finie) revient à admettre l'injectivité de cette application (l'espace peut donc s'identifier à un sous-espace du bidual)[1] . Avec les notations du théorème précédent, le rang de $(e_i)_{1 \leqslant i \leqslant p}$ est le rang de $(\widetilde{e}_i)_{1 \leqslant i \leqslant p}$, et le théorème (resp. ses corollaires) est un cas particulier du théorème **3-2**.1 (resp. des théorèmes **3-2**.8 et **3-2**.12)

3-3.2 Equations d'un s.e.v., d'un sous-espace affine

L'espace vectoriel $\mathbb{E}$ est de dimension n. Nous précisons ici le théorème **3-2**.9 dans cette situation particulière. Pour caractériser les vecteurs d'un sous-espace $\mathbb{F}$ de $\mathbb{E}$, de dimension p, nous obtenons deux moyens :

- Soit se donner une famille libre $(u_i)_{1 \leqslant i \leqslant p}$ de $\mathbb{F}$ (en fait une base de $\mathbb{F}$). Les vecteurs de $\mathbb{F}$ sont les combinaisons linéaires des $(u_i)_{1 \leqslant i \leqslant p}$.
- Soit se donner $n-p$ formes linéaires indépendantes $(\psi_i)_{1 \leqslant i \leqslant n-p}$ sur $\mathbb{E}$, identiquement nulles sur $\mathbb{F}$ (un "système d'équations linéaires" de $\mathbb{F}$). Les vecteurs de $\mathbb{F}$ sont exactement les vecteurs de $\mathbb{E}$ sur lesquels ces formes sont toutes nulles.

THÉORÈME 3-3.5 Si $\mathbb{F}$ est un sous-espace de dimension p d'un espace vectoriel $\mathbb{E}$ de dimension n, $\mathbb{F}$ est intersection de $n-p$ hyperplans indépendants. Plus précisément, il existe $n-p$ formes linéaires indépendantes $(\varphi_i)_{1 \leqslant i \leqslant n-p}$ sur $\mathbb{E}$ telles que

$$\forall x \in \mathbb{E} \qquad x \in \mathbb{F} \Longleftrightarrow \left\{ \begin{array}{c} \varphi_1(x) = 0 \\ \vdots \\ \varphi_{n-p}(x) = 0 \end{array} \right.$$

Tout système d'équations linéaires scalaires indépendantes caractérisant $\mathbb{F}$ est un système "équivalent" : si

$$\left\{ \begin{array}{c} \psi_1(x) = 0 \\ \vdots \\ \psi_r(x) = 0 \end{array} \right.$$

est un système d'équations définissant $\mathbb{F}$, avec les ψ_i formes linéaires indépendantes, alors

- $r = n - p$
- **Chaque ψ_j est combinaison linéaire des $(\varphi_i)_{1 \leqslant i \leqslant n-p}$.**
- **Chaque φ_j est combinaison linéaire des $(\psi_i)_{1 \leqslant i \leqslant n-p}$.**

Démonstration : Cet énoncé ne fait que paraphraser les résultats énoncés plus haut. Si l'on complète une base $(u_i)_{1\leqslant i\leqslant p}$ de $\mathbb{F}$ en une base $\mathcal{B} = (u_i)_{1\leqslant i\leqslant n}$ de $\mathbb{E}$, et qu'on note $\mathcal{B}^* = (\phi_i)_{1\leqslant i\leqslant n}$ la base duale, il est clair que

$$\mathbb{F} = \bigcap_{i=p+1}^{n} \ker \phi_i$$

Il suffit de poser, pour $i \in \{1, \dots, n-p\}$, $\varphi_i = \phi_{p+i}$ pour obtenir un système d'équations de $\mathbb{F}$. Si, à présent,

$$\begin{cases} \psi_1(x) = 0 \\ \quad\vdots \\ \psi_r(x) = 0 \end{cases} \quad \text{avec les } \psi_i \text{ formes linéaires indépendantes}$$

est un autre système d'équations de $\mathbb{F}$, la dimension de l'espace des solutions étant $n-r$ (théorème **3-2**.9) on a forcément $r = n - p$. Chaque ψ_j s'annule sur $\mathbb{F} = \bigcap_{i=1}^{n-p} \ker \varphi_i$ donc (théorème **3-2**.12) $\psi_j \in \operatorname{vect}(\varphi_i)_{1\leqslant i\leqslant n-p}$. De même chaque φ_j ... ■

Le résultat précédent possède évidemment un traduction "affine" :

COROLLAIRE 3-3.6 Si $\mathbb{A}$ est un sous-espace affine de dimension p d'un espace vectoriel $\mathbb{E}$ de dimension n, $\mathbb{A}$ est intersection de $n-p$ hyperplans affines indépendants. Plus précisément, il existe $n-p$ formes linéaires indépendantes $(\varphi_i)_{1\leqslant i\leqslant n-p}$ sur $\mathbb{E}$ et des scalaires $(b_i)_{1\leqslant i\leqslant n-p}$ telles que

$$\forall x \in \mathbb{E} \qquad x \in \mathbb{A} \Longleftrightarrow (\mathcal{S}) \begin{cases} \varphi_1(x) = b_1 \\ \quad\vdots \\ \varphi_{n-p}(x) = b_{n-p} \end{cases}$$

Tout autre système d'équations affines indépendantes définissant $\mathbb{A}$ est un système équivalent : ses équations sont combinaisons linéaires des équations de $(\mathcal{S})$ et réciproquement.

Démonstration : On dit bien sûr que des hyperplans affines sont indépendants si et seulement si leurs directions le sont. Il est clair ensuite que si φ est une forme linéaire non nulle sur $\mathbb{E}$ et b est un scalaire quelconque,

$$\varphi(x) = b$$

est l'équation d'un hyperplan affine de direction $\ker\varphi$. Si $\mathbb{A}$ passe par a et a pour direction $\mathbb{F}$ avec $\mathbb{F} = \bigcap_{i=1}^{n-p} \ker \varphi_i$ comme dans le théorème précédent, on a bien

$$x \in \mathbb{A} \Leftrightarrow x - a \in \mathbb{F} \Leftrightarrow (\mathcal{S}) \begin{cases} \varphi_1(x) = b_1 = \varphi_1(a) \\ \quad\vdots \\ \varphi_{n-p}(x) = b_{n-p} = \varphi_{n-p}(a) \end{cases} \qquad ■$$

Remarquons pour terminer que, réciproquement, si les formes $(\varphi_i)_{1\leqslant i\leqslant n-p}$ sont indépendantes, quels que soient les scalaires $(b_i)_{1\leqslant i\leqslant n-p}$, le système

$$(\mathcal{S})\left\{\begin{array}{c}\varphi_1(x)=b_1\\ \vdots\\ \varphi_{n-p}(x)=b_{n-p}\end{array}\right.$$

possède toujours au moins une solution (théorème **3-2.1**) et définit donc bien un sous-espace affine de direction $\bigcap_{i=1}^{n-p}\ker\varphi_i$, de dimension p (*cf.* section **2-1.7**).

3-4 Exercices

EXERCICE 3-4.1 Soit $E=\mathbb{R}_n[X]$ et $x_0<x_1<\cdots<x_n$ $n+1$ réels distincts.

1. Soit $\varphi_i\in E^*$ définie par $\varphi_i(P)=P(x_i)$. Montrer que $(\varphi_i)_{0\leqslant i\leqslant n}$ est une base de E^*.
2. Soit $(g_p)_{p\in\mathbb{N}}$ une suite d'éléments de E convergeant simplement sur $\mathbb{R}$ vers une fonction g. Montrer que g est un polynôme. Ce résultat est-il encore vrai avec $E=\mathbb{R}[X]$?

EXERCICE 3-4.2 Dans $(\mathbb{R}_2[X])^*$, trouver le rang de $\{\varphi_1,\varphi_2,\varphi_3\}$ avec

$$\varphi_1(P)=P'(a),\quad \varphi_2(P)=P'(b)\quad\text{et}\quad \varphi_3(P)=P'(c)\quad (a\neq b\neq c\neq a)$$

Même question dans $(\mathbb{R}_3[X])^*$ avec

$$\varphi_1(P)=P(a),\quad \varphi_2(P)=P(b),\quad \varphi_3(P)=P(c)\quad\text{et}\quad \varphi_4(P)=\int_a^b P(t)dt$$

EXERCICE 3-4.3 Soit E un K-ev et $\varphi,\psi\in E^*$ avec

$$\forall\ x\in E\quad \varphi(x).\psi(x)=0$$

Que peut-on dire de φ et ψ ?

EXERCICE 3-4.4 K est un corps commutatif et $n\in\mathbb{N}^*$:

1. Montrer que $\forall\ A,B,C\in M_n(K)$: $\mathrm{Trace}(ABC)=\mathrm{Trace}(BCA)$.
2. Si $\varphi\in(M_n(K))^*$, montrer que
$$\exists!\ A\in M_n(K)\quad\forall\ X\in M_n(K)\quad \varphi(X)=\mathrm{Trace}({}^tAX)$$
3. Si, de plus, $\forall\ A,B\quad \varphi(AB)=\varphi(BA)$, montrer que φ est proportionnelle à la trace.

EXERCICE 3-4.5 Etudier le rang de la famille de formes définies sur $\mathbb{R}^4$ par :

$$\left\{\begin{array}{ccccccccc}-x&+&by&+&cz&+&dt&=&\varphi_1(x,y,z,t)\\ ax&-&y&+&cz&+&dt&=&\varphi_2(x,y,z,t)\\ ax&+&by&-&z&+&dt&=&\varphi_3(x,y,z,t)\\ ax&+&by&+&cz&-&t&=&\varphi_4(x,y,z,t)\end{array}\right.$$

EXERCICE 3-4.6 Dans $E = C^\infty([-1,+1],\mathbf{R})$, soit $\varphi_k : f \longmapsto f^{(k)}(0)$ pour tout $k \in \mathbf{N}$. Montrer que la famille $(\varphi_k)_{k\geq 0}$ est libre dans E^*.

EXERCICE 3-4.7 Soient $(\alpha_i)_{1\leqslant i\leqslant n}$ et $(\beta_i)_{1\leqslant i\leqslant n}$ deux familles de complexes distincts deux à deux avec

$$\forall i\ \forall j \quad \alpha_i + \beta_j \neq 0$$

Résoudre le système d'équations

$$\forall i \in \{1,\dots,n\} \quad \sum_{j=1}^{n} \frac{x_j}{\alpha_i + \beta_j} = 1$$

EXERCICE 3-4.8 Soit E un sous-espace de dimension n de $\mathbf{R}^{\mathbf{R}}$. Pour $x \in \mathbf{R}$ on définit $\varepsilon_x \in E^*$ par

$$\varepsilon_x(f) = f(x)$$

1. Montrer que $(\varepsilon_x)_{x\in\mathbf{R}}$ est une famille génératrice de E^*.
2. Montrer qu'on peut trouver $(x_i)_{1\leqslant i\leqslant n} \in \mathbf{R}^n$ telle que $(\varepsilon_{x_i})_{1\leqslant i\leqslant n}$ soit une base de l'espace dual E^*.
3. Montrer qu'on peut trouver $(f_i)_{1\leqslant i\leqslant n} \in E^n$ avec $f_i(x_j) = 0$ pour $i \neq j$ et $f_i(x_i)=1$.
4. Soit f une fonction de classe C^∞ de $\mathbf{R}$ dans $\mathbf{R}$. On suppose que

$$\text{vect}\,(x \mapsto f(x+a))_{a\in\mathbf{R}}$$

est de dimension finie. Montrer que toutes les dérivées de f appartiennent à ce sous-espace. Que peut-on en déduire ?

EXERCICE 3-4.9 Soient λ et μ réels et a, b, c les racines complexes de

$$P(X) = X^3 - \lambda X + 1 - \mu$$

Discuter le système

$$\begin{cases} x + y + z = 1 \\ ax + by + cz = 2 \\ a^3x + b^3y + c^3z = 1 \end{cases}$$

EXERCICE 3-4.10 Suivant les valeurs du réel a, résoudre les systèmes :

$$\begin{cases} ax + y + z = 1 \\ x + ay + z = 1 \\ x + y + az = 1 \end{cases} \qquad \begin{cases} (a+1)x + y + z = a^2 + 3a \\ x + (a+1)y + z = a^3 + 3a^2 \\ x + y + (a+1)z = a^4 + 3a^3 \end{cases}$$

$$\begin{cases} 2x + 3y + z + 2t = 3 \\ 4x + 6y + 3z + 4t = 5 \\ 6x + 9y + 5z + 6t = 7 \\ 8x + 12y + 7z + at = 9 \end{cases}$$

EXERCICE 3-4.11 Résoudre dans $\mathbf{R}^n$:

$$\begin{cases} x_1 + x_2 + \cdots + x_n = 1 \\ x_1 + 2x_2 + \cdots + nx_n = 0 \\ \cdots \\ x_1 + 2^{n-1}x_2 + \cdots n^{n-1}x_n = 0 \end{cases} \qquad \begin{cases} ax_1 + x_2 + \cdots + x_n = 1 \\ x_1 + ax_2 + x_3 + \cdots + x_n = b \\ \cdots \\ x_1 + x_2 + x_3 + \cdots + ax_n = b^{n-1} \end{cases}$$

Chapitre 4

Calcul matriciel

4-1 Opérations sur les matrices. Matrice d'un morphisme

4-1.1 Espace $\mathcal{M}_{p,n}(\mathbb{K})$

DÉFINITION 4-1.1 *On appelle matrice à p lignes et n colonnes à coefficients dans $\mathbb{K}$ tout tableau rectangulaire de scalaires de $\mathbb{K}$ à p lignes et n colonnes :*

$$M = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$$

Plus précisément, un tel tableau est une famille de scalaires indexée par l'ensemble $\{1,\ldots,p\}\times\{1,\ldots,n\}$: par convention, le coefficient a_{ij} est situé à l'intersection de la $i^{ème}$ ligne et de la $j^{ème}$ colonne. On dit aussi que M est une matrice de type (p,n). L'ensemble de ces matrices est noté $\mathcal{M}_{p,n}(\mathbb{K})$. Lorsque $p=n$, on parle de matrices carrées, et on note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de type (n,n).

Un élément $\mathcal{M}_{p,n}(\mathbb{K})$ n'est pas autre chose qu'un np-uplet de scalaires, un élément de $\mathbb{K}^{np}$, avec ré-indexation des coordonnées. On en déduit évidemment le théorème :

THÉORÈME 4-1.2 $\mathcal{M}_{p,n}(\mathbb{K})$ **est un $\mathbb{K}$-ev de dimension np, pour les lois :**

$$(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} + (b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} = (a_{ij}+b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$$
$$\lambda\,(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} = (\lambda a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$$

De même la base canonique de $\mathbb{K}^{np}$ a une traduction matricielle :

COROLLAIRE 4-1.3 Pour $1 \leqslant i \leqslant p$ et $1 \leqslant j \leqslant n$, on note E_{ij}^{pn}, ou tout simplement E_{ij} s'il n'y a pas d'ambiguïté sur les valeurs de p et n, la matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ dont tous les coefficients sont nuls sauf celui de la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne qui vaut 1. La famille $(E_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ forme une base de $\mathcal{M}_{p,n}(\mathbb{K})$, qu'on appelle base canonique de $\mathcal{M}_{p,n}(\mathbb{K})$.

Dans cette base, on a simplement :

$$(a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} = \sum_{i=1}^{p}\sum_{j=1}^{n} a_{ij}E_{ij}$$

Il est aussi possible de se représenter de manière plus "compacte" un élément $M = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ de $\mathcal{M}_{p,n}(\mathbb{K})$. Deux visions sont possibles et "duales" l'une de l'autre : si on note

$$C_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{pj} \end{pmatrix} \in \mathbb{K}^p \quad \text{et} \quad L_i = (a_{i1}, \dots, a_{ij}, \dots, a_{in}) \in \mathbb{K}^n$$

respectivement la $j^{\text{ème}}$ colonne et la $i^{\text{ème}}$ ligne de M, on peut considérer que M est un n-uplet de vecteurs colonnes ou un p-uplet de vecteurs lignes :

$$M = (C_1, \dots, C_j, \dots, C_n) = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_p \end{pmatrix}$$

Cela revient en fait à identifier $\mathcal{M}_{p,n}(\mathbb{K})$ avec $(\mathbb{K}^p)^n$ ou $(\mathbb{K}^n)^p$.

4-1.2 Matrice associée à un morphisme

Soient $\mathbb{E}_n$ et $\mathbb{F}_p$ deux espaces vectoriels de dimensions respectives n et p. On en choisit des bases respectives $\mathcal{B} = (e_j)_{1\leqslant j\leqslant n}$ et $\mathcal{B}' = (e'_i)_{1\leqslant i\leqslant p}$.

DÉFINITION 4-1.4 *Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p)$, on appelle matrice de u dans les bases $\mathcal{B}$ et $\mathcal{B}'$, la matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ dont la $j^{\text{ème}}$ colonne est la colonne des coordonnées de $u(e_j)$ dans la base $\mathcal{B}'$. On note cette matrice*

$$M = M(u, \mathcal{B}, \mathcal{B}')$$

On notera que la matrice d'un morphisme se remplit et se lit (en général) par colonnes. Si $M(u, \mathcal{B}, \mathcal{B}') = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$, on a alors

$$\forall j \in \{1, \dots, n\} \quad u(e_j) = \sum_{i=1}^{p} a_{ij}e'_i$$

Lorsque u est un endomorphisme d'un espace $\mathbb{E}_n$, on travaille en général avec la même base au départ et à l'arrivée (ce n'est cependant pas toujours le cas, et il est parfois astucieux de bien choisir $\mathcal{B}' \neq \mathcal{B}$. Voir par exemple la section **4-2**). On note alors

$$M(u, \mathcal{B}) = M(u, \mathcal{B}, \mathcal{B}) \in \mathcal{M}_n(\mathbb{K})$$

et on parle de matrice de l'endomorphisme u dans la base $\mathcal{B}$.

Le théorème **2-3**.12 et la remarque qui le suit montrent qu'un morphisme entre espaces de dimensions finies est entièrement caractérisé par la donnée de sa matrice dans un couple de bases des espaces de départ et d'arrivée. Plus précisément :

THÉORÈME 4-1.5 Le choix d'une base $\mathcal{B}$ de $\mathbb{E}_n$ et d'une base $\mathcal{B}'$ de $\mathbb{F}_p$ détermine un isomorphisme

$$\varphi_{\mathcal{B},\mathcal{B}'} : u \mapsto M(u, \mathcal{B}, \mathcal{B}')$$

entre les espaces vectoriels $\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p)$ et $\mathcal{M}_{p,n}(\mathbb{K})$.

Démonstration : On vérifie en effet immédiatement que

$$M(u, \mathcal{B}, \mathcal{B}') + M(v, \mathcal{B}, \mathcal{B}') = M(u + v, \mathcal{B}, \mathcal{B}') \quad \text{et} \quad \lambda M(u, \mathcal{B}, \mathcal{B}') = M(\lambda u, \mathcal{B}, \mathcal{B}')$$

pour u et $v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p)$ et $\lambda \in \mathbb{K}$. Par le morphisme $\varphi_{\mathcal{B},\mathcal{B}'}$ la base canonique de $\mathcal{M}_{p,n}(\mathbb{K})$ est image de la base obtenue dans le théorème **2-3**.12. ∎

4-1.3 Coordonnées de l'image d'un vecteur

Les notations sont celles de la section précédente. On considère la matrice d'un morphisme $u \in \mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)$ comme un n-uplet de vecteurs colonnes :

$$M = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} = M(u, \mathcal{B}, \mathcal{B}') = (C_1, \ldots, C_j, \ldots, C_n)$$

et on se donne un vecteur x par la colonne de ses coordonnées dans $\mathcal{B}$:

$$\sum_{j=1}^{n} x_j e_j = x \longleftrightarrow X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

Son image par u est alors

$$y = u(x) = \sum_{j=1}^{n} x_j u(e_j)$$

La colonne Y des coordonnées de y dans $\mathcal{B}'$ est alors évidemment

$$Y = x_1 C_1 + \cdots + x_n C_n = \sum_{j=1}^{n} x_j C_j \in \mathbb{K}^p$$

On notera tout simplement

$$Y = MX$$

conformément à la règle opératoire suivante :

DÉFINITION 4-1.6 *Si $M \in \mathcal{M}_{p,n}(\mathbb{K})$, le produit MX de M par une* ***colonne*** *X appartenant à $\mathbb{K}^n$ (ou plus précisément à $\mathcal{M}_{n,1}(\mathbb{K})$, ayant donc autant de coordonnées que M a de* ***colonnes****) est la* ***colonne combinaison linéaire des colonnes*** *de M affectées des coefficients apparaissant dans X.*

Cette définition est conforme à celle donnée pour le produit d'une ligne $L \in \mathcal{M}_{1,n}(\mathbb{K})$ et d'une colonne $X \in \mathcal{M}_{n,1}(\mathbb{K})$ à la section **3-1.1**. On remarque d'ailleurs que, si M et X sont comme précédemment et si on considère cette fois M comme matrice des vecteurs lignes, on a aussi

$$Y = MX = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_p \end{pmatrix} X = \begin{pmatrix} L_1X \\ \vdots \\ L_iX \\ \vdots \\ L_pX \end{pmatrix}$$

4-1.4 Morphisme canoniquement associé à une matrice

Dans ce paragraphe, on identifie l'espace $\mathbb{K}^n$ à $\mathcal{M}_{n,1}(\mathbb{K})$ et on fait de même pour $\mathbb{K}^p$, ce qui revient à écrire les vecteurs de ces espaces sous forme de vecteurs colonnes.

DÉFINITION 4-1.7 *Si $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$, le morphisme canoniquement associé à A est l'application $\Phi_A \in \mathcal{L}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^p)$ définie par*

$$\Phi_A : \mathbb{K}^n \to \mathbb{K}^p \qquad X \mapsto AX$$

Il est évident que Φ_A est un morphisme. On a d'ailleurs clairement, si $\mathcal{B}_c$ et $\mathcal{B}'_c$ sont les bases canoniques respectives de $\mathbb{K}^n$ et $\mathbb{K}^p$:

$$A = M(\Phi_A, \mathcal{B}_c, \mathcal{B}'_c)$$

Le théorème **4-1**.5 s'écrit dans ce cas particulier

THÉORÈME 4-1.8 L'application $A \mapsto \Phi_A$ est un isomorphisme de $\mathcal{M}_{p,n}(\mathbb{K})$ dans $\mathcal{L}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^p)$.

Lorsqu'il n'y aura pas d'ambiguïté, il nous arrivera de "confondre" A et le morphisme Φ_A. On parlera alors du noyau de A (ensemble des vecteurs colonnes X de $\mathbb{K}^n$ vérifiant $AX = 0$) et de l'image de A (sous-espace de $\mathbb{K}^p$ engendré par les colonnes de la matrice A).

4-1.5 Produit matriciel

4-1.5.1 Matrice d'un composé de morphismes

On se donne trois espaces vectoriels $\mathbb{E}_n$, $\mathbb{F}_p$ et $\mathbb{G}_m$ rapportés à des bases respectives $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, $\mathcal{B}' = (e'_i)_{1 \leqslant i \leqslant p}$ et $\mathcal{B}'' = (e''_i)_{1 \leqslant i \leqslant m}$ et des morphismes

$$u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p) \text{ et } v \in \mathcal{L}_{\mathbb{K}}(\mathbb{F}_p, \mathbb{G}_m)$$

qu'on suppose déterminés par leurs matrices

$$A = M(u, \mathcal{B}, \mathcal{B}') \in \mathcal{M}_{p,n}(\mathbb{K}) \text{ et } B = M(v, \mathcal{B}', \mathcal{B}'') \in \mathcal{M}_{m,p}(\mathbb{K})$$

On est bien sûr cohérent dans le choix des bases, c'est-à-dire que la base de $\mathbb{F}_p$ considéré comme espace d'arrivée de u est aussi celle de l'espace de départ de v. **Le nombre de lignes de A est égal au nombre de colonnes de B.** On cherche à déterminer

$$C = M(v \circ u, \mathcal{B}, \mathcal{B}'') \in \mathcal{M}_{m,n}(\mathbb{K})$$

Pour ce faire, on écrit la matrice A sous forme de matrice de vecteurs colonnes et B sous forme de matrice de vecteurs lignes :

$$A = (C_1, \ldots, C_j, \ldots, C_n) \text{ et } B = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_m \end{pmatrix}$$

Si $j \in \{1, \ldots, n\}$ la $j^{\text{ème}}$ colonne de la matrice C représente les coordonnées de $v \circ u(e_j) = v[u(e_j)]$ dans $\mathcal{B}''$. Or les coordonnées de $u(e_j)$ dans $\mathcal{B}'$ sont, par définition de A, données par C_j. La $j^{\text{ème}}$ colonne de C est donc, conformément au résultat de la section **4-1.3**, égale à BC_j. Le coefficient de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne de la matrice C est donc égal au produit L_iC_j :

$$\text{si } A = (a_{kj})_{\substack{1 \leqslant k \leqslant p \\ 1 \leqslant j \leqslant n}} \text{ et } B = (b_{ik})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant k \leqslant p}} \text{ on a } C = (c_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \text{ avec } c_{ij} = \sum_{k=1}^{p} b_{ik} a_{kj}$$

Ce calcul amène à la définition du paragraphe suivant.

4-1.5.2 Produit matriciel

DÉFINITION 4-1.9 *Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$ on appelle produit BA la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ dont le coefficient de la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne est égal au produit de la $i^{\text{ème}}$ ligne de B par la $j^{\text{ème}}$ colonne de la matrice A.*

THÉORÈME 4-1.10 Avec les notations du paragraphe précédent, on a

$$M(v \circ u, \mathcal{B}, \mathcal{B}'') = M(v, \mathcal{B}', \mathcal{B}'')\, M(u, \mathcal{B}, \mathcal{B}')$$

COROLLAIRE 4-1.11 En particulier, en travaillant avec les morphismes canoniquement associés aux matrices, on a pour $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$

$$\Phi_B \circ \Phi_A = \Phi_{BA}$$

Le produit n'est évidemment défini que si **le nombre de lignes de A est égal au nombre de colonnes de B.** Le produit ainsi défini est parfois appelé "produit ligne-colonne" pour des raisons évidentes. Il est important d'avoir une vision un peu plus globale du produit matriciel, et ne pas se limiter à la formule

$$c_{ij} = L_i C_j = \sum_{k=1}^{p} b_{ik} a_{kj}$$

Le calcul effectué dans la section précédente montre que **la $j^{\text{ème}}$ colonne du produit BA est le produit de B par la $j^{\text{ème}}$ colonne de A, c'est-à-dire une combinaison linéaire des colonnes de B affectées des coefficients de la $j^{\text{ème}}$ colonne de A.**

Lorsque l'opération est possible, POST-MULTIPLIER une matrice B, c'est opérer des combinaisons sur les COLONNES de B en utilisant les coefficients des COLONNES de la matrice multiplicatrice.

4-1.5.3 Propriétés du produit matriciel

Le corollaire **4-1.11** montre que les propriétés du produit matriciel sont celles de la composition des morphismes : "associativité", distributivité à droite et à gauche par rapport à l'addition, exportativité des scalaires, non commutativité en général (lorsque les produits sont définis). Les résultats de la section **2-2.2** ont tous une traduction matricielle. En particulier :

THÉORÈME 4-1.12 L'application

$$\mathcal{M}_{m,p}(\mathbb{K}) \times \mathcal{M}_{p,n}(\mathbb{K}) \to \mathcal{M}_{m,n}(\mathbb{K}) \quad (B, A) \mapsto BA$$

est bilinéaire.

4-1.6 Transposition

Pour l'instant, les matrices ont plutôt été remplies par colonnes et les produits calculés en raisonnant sur les colonnes. La transposition permet de raisonner aussi sur les lignes.

DÉFINITION 4-1.13 *Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$, on appelle transposée de A et on note ${}^tA$ la matrice*

$${}^tA = (b_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} \in \mathcal{M}_{n,p}(\mathbb{K}) \ \text{ avec } b_{ij} = a_{ji}$$

obtenue en permutant les rôles joués par les indices de lignes et de colonnes.

Il est clair que la transposition réalise un isomorphisme d'espaces vectoriels entre $\mathcal{M}_{p,n}(\mathbb{K})$ et $\mathcal{M}_{n,p}(\mathbb{K})$. Si A est considérée comme matrice de vecteurs colonnes, ${}^tA$ est une matrice de vecteurs lignes et réciproquement :

$$A = (C_1, \dots, C_j, \dots, C_n) \text{ donne } {}^tA = \begin{pmatrix} L_1 = {}^tC_1 \\ \vdots \\ L_j = {}^tC_j \\ \vdots \\ L_n = {}^tC_n \end{pmatrix}$$

La transposition permute les facteurs d'un produit :

THÉORÈME 4-1.14 Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$, on a

$${}^t(BA) = {}^tA\,{}^tB$$

Démonstration : Si

$$A = (C_1, \dots, C_j, \dots, C_n) \text{ et } B = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_p \end{pmatrix}$$

le coefficient de la $j^{\text{ème}}$ ligne et de la $i^{\text{ème}}$ colonne de ${}^t(BA)$ vaut L_iC_j, celui de ${}^tA\,{}^tB$ vaut ${}^tC_j\,{}^tL_i$. Ils sont évidemment égaux. ■

On peut donc traduire sur les lignes les résultats énoncés précédemment sur les colonnes :

La $i^{\text{ème}}$ ligne du produit BA est le produit $i^{\text{ème}}$ ligne de B par A, c'est-à-dire une combinaison linéaire des lignes de A affectées des coefficients de la $i^{\text{ème}}$ ligne de B.

Lorsque l'opération est possible, PRE-MULTIPLIER une matrice A, c'est opérer des combinaisons sur les LIGNES de A en utilisant les coefficients des LIGNES de la matrice multiplicatrice.

4-1.7 Rang d'une matrice

DÉFINITION 4-1.15 *Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$, on appelle rang de A le rang de la famille de ses vecteurs colonnes dans $\mathbb{K}^p$.*

Le rang de A est donc le nombre maximal de colonnes indépendantes que l'on peut extraire de la matrice A. On a évidemment $\operatorname{rang} A \leqslant n$ et $\operatorname{rang} A \leqslant p$.
Si $A = M(u, \mathcal{B}, \mathcal{B}')$ (notations de la section **4-1.2**), les colonnes de A représentent les coordonnées des vecteurs de $u(\mathcal{B})$ dans $\mathcal{B}'$. Les relations de liaison éventuelles entre vecteurs de $u(\mathcal{B})$ se lisent donc sur ces colonnes, et on a

$$\operatorname{rang} M(u, \mathcal{B}, \mathcal{B}') = \operatorname{rang} u(\mathcal{B}) = \operatorname{rang} u$$

En particulier

$$\text{pour } A \in \mathcal{M}_{p,n}(\mathbb{K}) \qquad \text{rang}\, A = \text{rang}\, \Phi_A$$

Le rang d'un matrice peut également se lire sur ses lignes :

THÉORÈME 4-1.16 Pour $A \in \mathcal{M}_{p,n}(\mathbb{K})$ le rang de ${}^tA$ est égal au rang de la matrice A.

Démonstration : On détermine le rang de $\Phi_A : \mathbb{K}^n \to \mathbb{K}^p$ par application du théorème du rang. Si la matrice A est nulle, il n'y a rien à faire. Sinon on appelle r le rang du système des vecteurs lignes $(L_1, \dots, L_p)$ de A. Le noyau de A est formé des vecteurs X de $\mathbb{K}^n$ vérifiant le système d'équations linéaires :

$$\begin{cases} L_1 X = 0 \\ \quad\vdots \\ L_p X = 0 \end{cases}$$

La ligne L_i définit une forme linéaire ψ_i sur $\mathbb{K}^n$ et on recherche en fait $\bigcap_{i=1}^{p} \ker \psi_i$. Le théorème **3-2**.9 nous dit que cet espace est de dimension $n - r$. Par application du théorème du rang on obtient bien

$$\text{rang}\, \Phi_A = r \quad \blacksquare$$

Le théorème **2-3**.2 et la propriété $\Phi_B \circ \Phi_A = \Phi_{BA}$ donnent immédiatement :

THÉORÈME 4-1.17 Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{m,p}(\mathbb{K})$

$$\text{rang}\,(BA) \leqslant \inf(\text{rang}\, A, \text{rang}\, B)$$

On pourrait aussi remarquer que les colonnes de BA appartiennent toutes à l'espace engendré par les colonnes de B, et que les lignes de BA sont combinaisons linéaires des lignes de A.

4-1.8 Algèbre des matrices carrées $\mathcal{M}_n(\mathbb{K})$

4-1.8.1 Structure d'algèbre de $\mathcal{M}_n(\mathbb{K})$

La multiplication est une loi de composition interne dans $\mathcal{M}_n(\mathbb{K})$. Elle est associative, distributive à droite et à gauche par rapport à l'addition, vérifie

$$\forall \lambda \in \mathbb{K} \quad \forall A, B \in \mathcal{M}_n(\mathbb{K}) \qquad \lambda(AB) = (\lambda A)B = A(\lambda B)$$

Elle possède un élément neutre, la **matrice unité d'ordre** n (canoniquement associée à l'identité de $\mathbb{K}^n$) :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & & \vdots \\ \vdots & 0 & 1 & 0 & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

THÉORÈME 4-1.18 $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ **est une** $\mathbb{K}$**-algèbre. L'application définie par** $A \mapsto \Phi_A$ **est isomorphisme d'algèbres de** $\mathcal{M}_n(\mathbb{K})$ **dans** $\mathcal{L}_{\mathbb{K}}(\mathbb{K}^n)$**. Plus généralement, si** $\mathbb{E}_n$ **est un** $\mathbb{K}$**-ev de dimension** n**, le choix d'une base** $\mathcal{B}$ **de** $\mathbb{E}_n$ **détermine un isomorphisme**

$$u \mapsto M(u, \mathcal{B})$$

entre les algèbres $(\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n), +, \cdot, \circ)$ **et** $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$**.**

Les propriétés des algèbres $(\mathcal{L}_{\mathbb{K}}(\mathbb{K}^n), +, \cdot, \circ)$ et $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ sont donc identiques. En particulier, dès que $n \geqslant 2$, ces algèbres ne sont ni commutatives, ni intègres (*cf.* section **2-2.3**). Ces algèbres sont de dimension n^2, et si on note $(E_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ la base canonique de $\mathcal{M}_n(\mathbb{K})$ on a :

PROPOSITION 4-1.19 $E_{ij}E_{kl} = \delta_j^k E_{il}$

Démonstration : La vérification est immédiate : on post-multiplie E_{ij} par un matrice dont toutes les colonnes sont nulles, sauf la $l^{\text{ème}}$. La matrice produit a toutes ses colonnes nulles, sauf (peut-être) la $l^{\text{ème}}$ qui est la $k^{\text{ème}}$ colonne de E_{ij} (les coefficients de la $l^{\text{ème}}$ colonne de E_{kl} sont tous nuls, sauf un 1 à la $k^{\text{ème}}$ ligne). Le produit vaut donc 0_n si $j \neq k$ et E_{il} sinon. ■

EXERCICE 4-1.20 Montrer que le centre de l'algèbre $\mathcal{M}_n(\mathbb{K})$ (*cf.* **2-2.4**) est formé de la droite vectorielle $\mathbb{K} \cdot I_n$ des matrices dites "scalaires". On retrouve ainsi le résultat de **2-2.4** dans le cas de la dimension finie. Indication : pour qu'une matrice

$$M = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} E_{ij}$$

commute avec toute matrice de $\mathcal{M}_n(\mathbb{K})$, il faut et il suffit qu'elle commute avec toute matrice d'une base de $\mathcal{M}_n(\mathbb{K})$.

4-1.8.2 Groupe linéaire $\mathcal{GL}_n(\mathbb{K})$

On traduit ici matriciellement les résultats des sections **2-2.5** et **2-3.2** :

THÉORÈME 4-1.21 L'ensemble des matrices carrées d'ordre n **inversibles pour la multiplication forme un groupe multiplicatif qu'on appelle groupe linéaire d'ordre** n **du corps** $\mathbb{K}$**. Ce groupe** $(\mathcal{GL}_n(\mathbb{K}), \times)$ **est non commutatif dès que** $n \geqslant 2$**. Le choix d'une base** $\mathcal{B}$ **d'un** $\mathbb{K}$**-ev** $\mathbb{E}_n$ **de dimension** n **détermine un isomorphisme de groupes** $u \mapsto M(u, \mathcal{B})$ **de** $\mathcal{GL}_{\mathbb{K}}(\mathbb{E}_n)$ **vers** $\mathcal{GL}_n(\mathbb{K})$**.**

THÉORÈME 4-1.22 Si $A \in \mathcal{M}_n(\mathbb{K})$**, les propriétés suivantes sont équivalentes :**

- $A \in \mathcal{GL}_n(\mathbb{K})$
- **Le rang de** A **est égal à** n
- **Les** n **colonnes de** A **sont indépendantes**

- **Les n lignes de A sont indépendantes**
- **A est inversible à droite pour $\times$ dans $\mathcal{M}_n(\mathbb{K})$**
- **A est inversible à gauche pour $\times$ dans $\mathcal{M}_n(\mathbb{K})$**
- **A est régulière à droite pour $\times$ dans $\mathcal{M}_n(\mathbb{K})$**
- **A est régulière à gauche pour $\times$ dans $\mathcal{M}_n(\mathbb{K})$**
- **${}^tA$ est inversible (et on a alors $({}^tA)^{-1} = {}^t(A^{-1})$)**
- **La seule solution dans $\mathbb{K}^n$ de l'équation $AX = 0$ est la colonne nulle. (résoudre $AX = 0$ revient d'ailleurs à chercher les relations de liaison entre les colonnes de A)**

La seule affirmation à démontrer est peut-être $({}^tA)^{-1} = {}^t(A^{-1})$. Elle découle immédiatement du calcul :

$${}^t(AA^{-1}) = {}^tI_n = I_n = {}^t(A^{-1})\,{}^tA$$

THÉORÈME 4-1.23 Si u est un isomorphisme de $\mathbb{E}_n$ vers $\mathbb{E}'_n$, et si A est égale à $M(u, \mathcal{B}, \mathcal{B}')$, cette matrice est inversible et $A^{-1} = M(u^{-1}, \mathcal{B}', \mathcal{B})$. En particulier, si $u \in \mathcal{GL}(\mathbb{E}_n)$

$$A = M(u, \mathcal{B}) \Longrightarrow A^{-1} = M(u^{-1}, \mathcal{B})$$

Ce théorème résulte immédiatement du théorème **4-1**.10.

THÉORÈME 4-1.24 On ne change pas le rang d'une matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ en la prémultipliant par une matrice de $\mathcal{GL}_p(\mathbb{K})$ ou en la postmultipliant par une matrice de $\mathcal{GL}_n(\mathbb{K})$.

EXERCICE 4-1.25 Déterminer le centre du groupe $(\mathcal{GL}_n(\mathbb{K}), \times)$.

4-1.8.3 Sous algèbres remarquables de $\mathcal{M}_n(\mathbb{K})$

- **Matrices scalaires :**

  Nous avons déja rencontré la sous-algèbre $\mathbb{K}\cdot I_n$ des matrices scalaires qui ont la particularité de commuter avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$.

- **Matrices diagonales :**

  On appelle **diagonale principale** d'une matrice carrée $M = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ la diagonale des coefficients $(a_{ii})_{1 \leqslant i \leqslant n}$. Une matrice carrée est dite diagonale si et seulement si les coefficients hors de la diagonale principale sont tous nuls. En travaillant dans la base canonique de $\mathcal{M}_n(\mathbb{K})$, l'ensemble des matrices diagonales est exactement $\mathrm{vect}(E_{ii})_{1 \leqslant i \leqslant n}$ et on note

$$\mathrm{Diag}(a_1, \ldots, a_n) = \sum_{i=1}^{n} a_i E_{ii} = \begin{pmatrix} a_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & & \vdots \\ \vdots & 0 & a_i & 0 & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & a_n \end{pmatrix}$$

On a clairement

$$\operatorname{Diag}(a_1, \dots, a_n) \times \operatorname{Diag}(b_1, \dots, b_n) = \operatorname{Diag}(a_1 b_1, \dots, a_n b_n)$$

et on en déduit que les matrices diagonales forment une sous-algèbre commutative $\mathcal{D}_n$ de $\mathcal{M}_n(\mathbb{K})$. Il est de plus clair qu'une matrice diagonale $\operatorname{Diag}(a_1, \dots, a_n)$ est inversible si et seulement si tous les a_i sont non nuls et on a alors

$$\operatorname{Diag}(a_1, \dots, a_n)^{-1} = \operatorname{Diag}(a_1^{-1}, \dots, a_n^{-1})$$

- **Matrices triangulaires supérieures.**

On dit qu'une matrice carrée $M = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ est triangulaire supérieure si et seulement si, pour $i > j$, on a $a_{ij} = 0$, ce qui signifie que les coefficients sous la diagonale principale sont tous nuls. La matrice est dite triangulaire supérieure stricte si $a_{ij} = 0$ est vérifié pour tous i et j avec $i \geqslant j$. L'ensemble $\mathcal{T}_n$ des matices triangulaires supérieures est donc le sous-espace de $\mathcal{M}_n(\mathbb{K})$

$$\mathcal{T}_n = \operatorname{vect}(E_{ij})_{1 \leqslant i \leqslant j \leqslant n}$$

Il est de dimension $\dfrac{n(n+1)}{2}$. $\mathcal{T}_n$ est stable pour le produit : si A et B sont deux matrices triangulaires supérieures, la $j^{\text{ème}}$ colonne du produit AB est combinaison linéaire des j premières colonnes de A (car les coefficients de la $j^{\text{ème}}$ colonne de B sont nuls à partir de l'indice $j+1$). Comme ces colonnes sont toutes dans le sous-espace $\operatorname{vect}(e_1, \dots, e_j)$ de $\mathbb{K}^n$ rapporté à sa base canonique, il en résulte clairement que AB est triangulaire supérieure. $\mathcal{T}_n$ est donc une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$. De plus, le produit de la $i^{\text{ème}}$ ligne de A et de la $i^{\text{ème}}$ colonne de B est égal au produit $a_{ii}b_{ii}$ des éléments diagonaux.

La caractérisation géométrique des endomorphismes d'un espace de dimension n possèdant, dans une base $\mathcal{B}$, une matrice triangulaire supérieure est simple. Commençons par la définition :

DÉFINITION 4-1.26 *Si $\mathbb{E}_n$ est un espace vectoriel de dimension n, on appelle drapeau de sous-espaces vectoriels de $\mathbb{E}_n$ toute famille $(\mathbb{E}_1, \dots, \mathbb{E}_n)$ de sous-espaces de $\mathbb{E}_n$, avec*

$$\mathbb{E}_1 \subsetneq \mathbb{E}_2 \subsetneq \cdots \subsetneq \mathbb{E}_n$$

Il est alors clair que $\mathbb{E}_i$ est de dimension i. Une **base de $\mathbb{E}_n$ adaptée à ce drapeau** est une base $(e_i)_{1 \leqslant i \leqslant n}$ avec

$$\forall i \qquad \mathbb{E}_i = \operatorname{vect}(e_1, \dots, e_i)$$

Il est clair qu'inversement, toute base de $\mathbb{E}_n$ permet de définir un unique drapeau qui lui soit adapté. La notion de drapeau est alors liée à celle de matrice triangulaire supérieure par la proposition (évidente pour qui a compris la définition de la matrice d'un endomorphisme) :

PROPOSITION 4-1.27 Si u est un endomorphisme d'un espace $\mathbb{E}_n$ de dimension n et si $\mathcal{B}$ est une base de $\mathbb{E}_n$, la matrice de u dans la base $\mathcal{B}$ est triangulaire supérieure ssi les sous-espaces du drapeau associé à $\mathcal{B}$ sont tous stables par u :

$$\forall i \qquad u\left(\text{vect}\left(e_1, \ldots, e_i\right)\right) \subset \text{vect}\left(e_1, \ldots, e_i\right)$$

On peut bien sûr en déduire une autre démonstration du fait que l'ensemble des matrices triangulaires supérieures forment une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$. Il suffit de montrer que, $\mathbb{E}_n$ et $\mathcal{B}$ étant fixés, l'ensemble des endomorphismes de $\mathbb{E}_n$ vérifiant la condition précédente est une sous-algèbre de $\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$.

EXERCICE 4-1.28 Avec les notations précédentes, montrer que la matrice de u est triangulaire supérieure stricte ssi

$$u(e_1) = 0_{\mathbb{E}_n} \quad \text{et} \quad \forall i \geqslant 2 \quad u\left(\text{vect}\left(e_1, \ldots, e_i\right)\right) \subset \text{vect}\left(e_1, \ldots, e_{i-1}\right)$$

En déduire qu'alors $u^n = 0_{\mathcal{L}(\mathbb{E}_n)}$.

Si la matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ de u dans la base $\mathcal{B}$ est triangulaire supérieure, dire que $a_{ii} = 0$ revient à écrire $u\left(\text{vect}\left(e_1, \ldots, e_i\right)\right) \subset \text{vect}\left(e_1, \ldots, e_{i-1}\right)$. Comme un automorphisme d'un espace vectoriel conserve les dimensions des sous-espaces, il est clair que :

PROPOSITION 4-1.29 Une matrice triangulaire supérieure T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls . Son inverse est alors triangulaire supérieure et les coefficients diagonaux de T^{-1} sont alors les inverses des coefficients diagonaux de T.

La dernière affirmation provient du fait que, dans un produit de matrices de $\mathcal{T}_n$, les coefficients diagonaux se calculent très facilement. Le fait que l'inverse d'une matrice de $\mathcal{T}_n \cap \mathcal{GL}_n(\mathbb{K})$ est dans $\mathcal{T}_n$ (et la même propriété pour les matrices diagonales) est un cas particulier de l'exercice suivant :

EXERCICE 4-1.30 Si $\mathbb{A}$ est une $\mathbb{K}$-algèbre et $\mathbb{B}$ en est une sous-algèbre de dimension finie, montrer que tout élément de $\mathbb{B}$ inversible dans $\mathbb{A}$ a son inverse dans $\mathbb{B}$. (Indication : considérer l'application $\mathbb{B} \rightarrow \mathbb{B} \quad x \mapsto bx$)

4-1.8.4 Caractérisation du rang à l'aide de matrices extraites

DÉFINITION 4-1.31 *Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$ et p' et n' sont deux entiers vérifiant $1 \leqslant p' \leqslant p$ et $1 \leqslant n' \leqslant n$, on appelle matrice de type (p', n') extraite de A toute matrice de $\mathcal{M}_{p',n'}(\mathbb{K})$ obtenue en rayant $p - p'$ lignes et $n - n'$ colonnes de A.*

Le résultat qui suit est une caractérisation **théorique** du rang d'une matrice. Dans la pratique, il est extrêmement rare qu'on l'utilise pour calculer effectivement le rang d'une matrice.

THÉORÈME 4-1.32 Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$ une matrice non nulle. Les propriétés suivantes sont équivalentes :

i) **Le rang de A est égal à r.**

ii) **Il existe une matrice carrée d'ordre r extraite de A inversible, toute matrice extraite d'ordre $> r$ (s'il en existe) étant non inversible.**

iii) **Il existe une matrice carrée d'ordre r extraite de A inversible, toute matrice extraite d'ordre $r+1$ (s'il en existe) étant non inversible.**

Démonstration : Si A est de rang r, elle possède r colonnes indépendantes. La matrice $A' \in \mathcal{M}_{p,r}(\mathbb{K})$ extraite en ne conservant que ces r colonnes est encore de rang r et possède donc r lignes indépendantes. La matrice A'' extraite de A' en ne conservant que ces r lignes est donc dans $\mathcal{GL}_r(\mathbb{K})$. De plus, si B est une matrice carrée d'ordre $q > r$, les q colonnes de A dont elle provient sont nécessairement liées (puisque rang $A = r$) et une relation de liaison entre ces colonnes donne *a fortiori* une relation de liaison entre les colonnes de B. B est donc non inversible. On a prouvé $i) \Rightarrow ii)$. Il reste à prouver $iii) \Rightarrow i)$. S'il existe une matrice carrée d'ordre r inversible extraite de A, les r colonnes de A dont elle provient sont indépendantes donc rang $A \geqslant r$. Si on avait rang $A > r$, on pourrait extraire $r+1$ colonnes indépendantes de A et donc, comme précédemment, une matrice carrée d'ordre $r+1$ inversible. ■

Le rang de A est donc la taille maximale d'une matrice carrée inversible que l'on peut extraire de A.

4-2 Changements de bases

4-2.1 Matrice de passage

DÉFINITION 4-2.1 *Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est une base de $\mathbb{E}_n$ et si $\mathcal{F} = (u_j)_{1 \leqslant j \leqslant p}$ est une famille finie de vecteurs de $\mathbb{E}_n$, on appelle matrice représentative de $\mathcal{F}$ dans la base $\mathcal{B}$ la matrice $M = M_{\mathcal{B}}(\mathcal{F}) \in \mathcal{M}_{n,p}(\mathbb{K})$ dont la $j^{\text{ème}}$ colonne est formée des coordonnées du vecteur u_j dans $\mathcal{B}$. Le rang de cette matrice est donc évidemment égal au rang de $\mathcal{F}$.*

DÉFINITION 4-2.2 *Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ et $\mathcal{B}' = (e'_i)_{1 \leqslant i \leqslant n}$ sont deux bases d'un espace vectoriel $\mathbb{E}_n$, on appelle matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ la matrice représentative des vecteurs de $\mathcal{B}'$ dans $\mathcal{B}$.*

Cette matrice carrée est inversible puisque ses colonnes sont indépendantes. **Elle se remplit** (en général) **par colonnes**, chaque colonne représentant un vecteur de $\mathcal{B}'$ (la "nouvelle base") exprimé par ses coordonnées dans $\mathcal{B}$ ("l'ancienne base"). Si

$$P = M_{\mathcal{B}}(\mathcal{B}')$$

est cette matrice, la définition de la matrice d'une application linéaire par rapport à un couple de bases permet aussi d'écrire

$$P = M_{\mathcal{B}}(\mathcal{B}') = M(id_{\mathbb{E}}, \mathcal{B}', \mathcal{B}) \quad \text{attention à l'ordre des bases!}$$

THÉORÈME 4-2.3 Si $P \in \mathcal{GL}_n(\mathbb{K})$ est la matrice de passage d'une base $\mathcal{B}$ à une base $\mathcal{B}'$ et Q est la matrice de passage de $\mathcal{B}'$ à une base $\mathcal{B}''$ on a :

- **La matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$ est égale à P^{-1}.**
- **La matrice de passage de $\mathcal{B}$ à $\mathcal{B}''$ est égale au produit PQ.**

Démonstration : On a en effet :

$$M(id_{\mathbb{E}_n}, \mathcal{B}', \mathcal{B})^{-1} = M(id_{\mathbb{E}_n}^{-1}, \mathcal{B}, \mathcal{B}') = M(id_{\mathbb{E}_n}, \mathcal{B}, \mathcal{B}')$$

$$M(id_{\mathbb{E}_n}, \mathcal{B}'', \mathcal{B}) = M(id_{\mathbb{E}_n}, \mathcal{B}', \mathcal{B}) \times M(id_{\mathbb{E}_n}, \mathcal{B}'', \mathcal{B}')$$

puisque $id_{\mathbb{E}_n} = id_{\mathbb{E}_n} \circ id_{\mathbb{E}_n}$. ■

Attention à l'ordre dans lequel on multiplie les matrices de passage ! On multiplie de la gauche vers la droite au fur et à mesure des changements. C'est l'inverse de ce qu'on fait lorsqu'on calcule la matrice d'un composé de morphismes, quand la multiplication des matrices se lit plutôt de droite à gauche.

REMARQUE 4-2.4 Il est à noter que, réciproquement, toute matrice de $\mathcal{GL}_n(\mathbb{K})$ peut être considérée comme une matrice de passage (par exemple de la base canonique de $\mathbb{K}^n$ à la base de ses vecteurs colonnes). Une matrice inversible gagne parfois à être considérée sous cet angle, plutôt que comme matrice d'un automorphisme. Voir notamment ce qui sera dit plus loin sur les matrices orthogonales (chapitre sur les espaces euclidiens).

4-2.2 Changement de coordonnées d'un vecteur

THÉORÈME 4-2.5 Un vecteur $x \in \mathbb{E}_n$ peut être représenté par le vecteur colonne X de ses coordonnées par rapport à $\mathcal{B}$ et par la colonne X' des ses coordonnées dans $\mathcal{B}'$. Si $P = M_{\mathcal{B}}(\mathcal{B}')$, on a

$$X = PX'$$

Démonstration : On traduit matriciellement l'égalité

$$x = id_{\mathbb{E}_n}(x)$$

comme on l'a vu au paragraphe **4-1.3**, en travaillant avec la base $\mathcal{B}'$ au départ et $\mathcal{B}$ à l'arrivée. ■

Se souvenir que cette formule donne les "anciennes" coordonnées en fonction des "nouvelles". Pour trouver X', on est contraint à résoudre le système d'inconnue X', dont la solution théorique est bien évidemment

$$X' = P^{-1}X$$

La formule de changement de base est donc une "mauvaise" formule pour déterminer effectivement les coordonnées d'un vecteur. C'est au contraire une "bonne" formule quand il s'agit de transformer un problème par "changement de variables linéaire" : une expression $\Phi(X)$ dépendant des coordonnées de x dans $\mathcal{B}$ s'exprime $\Phi(PX')$ si on décide de travailler dans $\mathcal{B}'$.

EXERCICE 4-2.6 (Changement de bases duales) Montrer que, si P est matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, la matrice de passage de $\mathcal{B}^*$ à $(\mathcal{B}')^*$ est égale à ${}^tP^{-1}$. (Remarquer que la relation $X = PX'$ donne une expression des formes coordonnées dans $\mathcal{B}$ en fonctions des formes coordonnées dans $\mathcal{B}'$).

4-2.3 Changement de la matrice d'un morphisme. Matrices équivalentes

THÉORÈME 4-2.7 Soit un morphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p)$ caractérisé par ses matrices

$$A = M(u, \mathcal{B}_1, \mathcal{B}'_1) \quad \textbf{et} \quad B = M(u, \mathcal{B}_2, \mathcal{B}'_2)$$

par rapport à des couples de bases $(\mathcal{B}_1, \mathcal{B}'_1)$ et $(\mathcal{B}_2, \mathcal{B}'_2)$. On note

$$P = M_{\mathcal{B}_1}(\mathcal{B}_2) \in \mathcal{GL}_n(\mathbb{K}) \text{ et } Q = M_{\mathcal{B}'_1}(\mathcal{B}'_2) \in \mathcal{GL}_p(\mathbb{K})$$

les matrices de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$ et de $\mathcal{B}'_1$ à $\mathcal{B}'_2$. On a alors

$$B = Q^{-1}AP$$

Démonstration : L'égalité $u = id_{\mathbb{F}_p} \circ u \circ id_{\mathbb{E}_n}$ possède la traduction matricielle :

$$M(u, \mathcal{B}_2, \mathcal{B}'_2) = M\left(id_{\mathbb{F}_p}, \mathcal{B}'_1, \mathcal{B}'_2\right) \times M(u, \mathcal{B}_1, \mathcal{B}'_1) \times M(id_{\mathbb{E}_n}, \mathcal{B}_2, \mathcal{B}_1)$$

ce qui donne le résultat escompté. ■

Ceci amène à la définition suivante :

DÉFINITION 4-2.8 *Si A et $B \in \mathcal{M}_{p,n}(\mathbb{K})$, on dit que B est équivalente à A si et seulement si il existe un couple de matrices inversibles*

$$\exists (M, N) \in \mathcal{GL}_p(\mathbb{K}) \times \mathcal{GL}_n(\mathbb{K}) \quad B = MAN$$

Comme une matrice inversible peut toujours être interprétée comme matrice de passage, il est équivalent de dire que A et B représentent le même morphisme $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^p)$ (ou plus généralement d'un $\mathbb{K}$-ev de dimension n vers un $\mathbb{K}$-ev de dimension p) pour des couples de bases (en général) différents.

Il s'agit clairement d'une relation d'équivalence dans $\mathcal{M}_{p,n}(\mathbb{K})$. Le problème de la classification des matrices de $\mathcal{M}_{p,n}(\mathbb{K})$ à équivalence près est entièrement résolu par le théorème suivant :

THÉORÈME 4-2.9 Deux matrices de $\mathcal{M}_{p,n}(\mathbb{K})$ sont équivalentes si et seulement si elles ont même rang. Plus précisément, une matrice de type (p,n) et de rang $r > 0$ est équivalente à la matrice "canonique" de type (p,n) et de rang r

$$J_r^{n,p} = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{p-r,r} & 0_{p-r,n-r} \end{pmatrix}$$

(notée simplement J_r s'il n'y a pas d'ambiguïté).

Démonstration : Il est clair que deux matrices équivalentes ont même rang (à cause de l'interprétation géométrique de l'équivalence, ou à cause du théorème **4-1**.24). Réciproquement, si on montre que toute matrice de rang r est équivalente à J_r, on aura prouvé, par transitivité, que deux matrices de même rang (et de même type!) sont équivalentes. Nous donnons ici une démonstration "géométrique", et reverrons sans doute une démonstration plus "algorithmique" dans la section relative aux manipulations élémentaires. Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$ de rang r. On peut interpréter A comme matrice, par rapport aux bases canoniques, du morphisme $\Phi_A \in \mathcal{L}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^p)$ canoniquement associé. La recherche d'un couple de bases où la matrice de Φ_A soit égale à J_r n'est pas difficile (d'un point de vue théorique) si on regarde la matrice J_r attentivement :

Au **départ**, les $n - r$ dernières colonnes de J_r étant nulles, on choisit une base $\mathcal{B}$ de $\mathbb{K}^n$ adaptée à la décomposition

$$\mathbb{K}^n = \mathbb{S}_r \oplus \ker \Phi_A$$

où $\mathbb{S}_r$ est un supplémentaire (de dimension r : théorème du rang) de $\ker \Phi_A$. Notons $(u_j)_{1 \leqslant j \leqslant r}$ les r premiers vecteurs de cette base.

A **l'arrivée**, il n'y a pas le choix pour les r premiers vecteurs de la base de $\mathbb{K}^p$: ce sont les vecteurs $v_j = \Phi_A(u_j)$ (voir les r premières colonnes de J_r). Ces vecteurs formant un système libre (théorème fondamental d'isomorphisme **2-1**.36) peuvent effectivement être complétés en une base $\mathcal{B}'$ de $\mathbb{K}^p$. On a alors clairement

$$J_r = M(\Phi_A, \mathcal{B}, \mathcal{B}') \quad \blacksquare$$

Il y a donc dans $\mathcal{M}_{p,n}(\mathbb{K})$ exactement $\inf(p,n) + 1$ classes d'équivalence pour la relation ... d'équivalence.

EXERCICE 4-2.10 Soient A et $B \in \mathcal{M}_{p,n}(\mathbb{K})$ tels qu'il existe des matrices carrées P, P', Q et Q' telles que $B = QAP$ et $A = Q'BP'$. Montrer que A et B sont équivalentes.

EXERCICE 4-2.11 Retrouver ainsi le fait qu'une matrice et sa transposée ont toujours même rang.

4-2.4 Changement de la matrice d'un endomorphisme. Matrices semblables

On applique ici les résultats du paragraphe précédent lorsqu'on travaille avec les matrices d'un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$, donc avec les mêmes bases au départ et à l'arrivée.

THÉORÈME 4-2.12 Soit un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$ caractérisé par ses matrices

$$A = M(u, \mathcal{B}_1) \quad \textbf{et} \quad B = M(u, \mathcal{B}_2)$$

par rapport à des bases respectives $\mathcal{B}_1$ et $\mathcal{B}_2$. On note

$$P = M_{\mathcal{B}_1}(\mathcal{B}_2) \in \mathcal{GL}_n(\mathbb{K})$$

la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. On a alors

$$B = P^{-1}AP$$

DÉFINITION 4-2.13 *Si A et $B \in \mathcal{M}_n(\mathbb{K})$, on dit que B est semblable à A si et seulement si il existe une matrice inversible $P \in \mathcal{GL}_n(\mathbb{K})$ telle que*

$$B = P^{-1}AP$$

Il est équivalent de dire que A et B représentent le même morphisme $u \in \mathcal{L}(\mathbb{K}^n)$ (ou plus généralement d'un $\mathbb{K}$-ev de dimension n) dans des bases (en général) différentes.

Là encore, il s'agit clairement d'une relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$. Une classe d'équivalence est appelée classe de similitude. La classification des matrices de $\mathcal{M}_n(\mathbb{K})$ à une similitude près, c'est-à-dire la recherche des différentes classes de similitude, est un problème plus compliqué que pour la relation d'équivalence. Ce sera en partie l'objet du chapitre sur la réduction des endomorphismes, et nous n'en donnerons qu'une solution partielle.

S'il est clair que deux matrices semblables ont même rang, la réciproque est fausse : par exemple dans $\mathcal{M}_2(\mathbb{K})$, les matrices $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ sont de rang 2 mais ne sont pas semblables, puisque la classe de similitude de I_2 est clairement réduite à $\{I_2\}$.

EXERCICE 4-2.14 Déterminer dans $\mathcal{M}_n(\mathbb{K})$ les classes de similitude réduites à 1 élément. (il y a au moins les classes des matrices scalaires, parce que notamment la matrice de l'homothétie λid_{E_n} est, **dans toute base** de $\mathbb{E}_n$, égale à λI_n). Voir les exercices **4-1**.20 et **4-1**.25.

DÉFINITION 4-2.15 *On appelle invariant de similitude dans $\mathcal{M}_n(\mathbb{K})$ toute application de $\mathcal{M}_n(\mathbb{K})$ dans un ensemble $\mathbb{X}$ quelconque, constante sur chaque classe de similitude.*

Le rang est, par exemple, un tel invariant. Pour montrer que deux matrices de $\mathcal{M}_n(\mathbb{K})$ ne sont pas semblables, il suffit parfois de trouver un invariant de similitude prenant des valeurs distinctes sur ces deux matrices. C'est ainsi qu'on peut dire au premier coup d'oeil que les matrices de $\mathcal{M}_3(\mathbf{R})$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

ne sont pas semblables. C'est ici le rang qui permet de les départager. Nous verrons ultérieurement que le déterminant, le polynôme caractéristique sont des invariants de similitude. Il est clair en tout cas qu'un invariant de similitude sur $\mathcal{M}_n(\mathbb{K})$

$$s : \mathcal{M}_n(\mathbb{K}) \to \mathbb{X}$$

permet de définir une application (qu'on notera encore s par abus d'écriture)

$$s : \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n) \to \mathbb{X}$$

où, pour u endomorphisme de $\mathbb{E}_n$, $s(u)$ sera simplement défini comme étant l'image par s d'une matrice représentative de u dans une base arbitraire de $\mathbb{E}_n$ (puisque le résultat de l'opération ne dépend pas de la base choisie). C'est ainsi que le rang d'un endomorphisme est lié au rang d'une de ses matrices.

Un autre invariant simple est la trace :

DÉFINITION 4-2.16 *Si $A \in \mathcal{M}_n(\mathbb{K})$, on appelle trace de A la somme des coefficients de la diagonale principale de A :*

$$A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \Rightarrow \operatorname{trace}(A) = \sum_{i=1}^{n} a_{ii}$$

La trace est évidemment une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$.

THÉORÈME 4-2.17 Si A et $B \in \mathcal{M}_n(\mathbb{K})$ on a

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

Démonstration : Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ et $B = (b_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$, le coefficient diagonal d'indice i du produit AB vaut

$$c_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}$$

et donc

$$\operatorname{trace}(AB) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} a_{ik} = \operatorname{trace}(BA) \quad \blacksquare$$

On a en fait un résultat analogue si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$ de manière à ce que les produits AB et BA soient définis et appartiennent respectivement à $\mathcal{M}_p(\mathbb{K})$ et $\mathcal{M}_n(\mathbb{K})$.

COROLLAIRE 4-2.18 La trace est un invariant de similitude sur $\mathcal{M}_n(\mathbb{K})$.

Démonstration : Si $P \in \mathcal{GL}_n(\mathbb{K})$ et $M \in \mathcal{M}_n(\mathbb{K})$, on a en effet

$$\begin{aligned}\operatorname{trace}\left(P^{-1}MP\right) &= \operatorname{trace}\left(P^{-1}(MP)\right)\\ &= \operatorname{trace}\left((MP)P^{-1}\right) = \operatorname{trace}(M) \quad \blacksquare\end{aligned}$$

On peut donc, conformément à ce qui a été vu plus haut, définir la trace d'un endomorphisme d'un espace vectoriel de dimension finie :

DÉFINITION 4-2.19 *Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$, on appelle trace de u la trace d'une matrice représentative de u dans une base quelconque de $\mathbb{E}_n$. La trace est une forme linéaire sue $\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$ et on a :*

$$\forall u, v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n) \quad \operatorname{trace}(uv) = \operatorname{trace}(vu)$$

REMARQUE 4-2.20 Si p est un projecteur sur un sous-espace $\mathbb{E}_r$ parallèlement à un supplémentaire $\mathbb{E}_{n-r}$, l'écriture de la matrice de p dans une base adaptée à la décomposition $\mathbb{E} = \mathbb{E}_r \oplus \mathbb{E}_{n-r}$ est très simple :

$$M(p, \mathcal{B}) = \begin{pmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{pmatrix}$$

et on obtient $\operatorname{trace}(p) = r.1_{\mathbb{K}}$. En particulier, si $\mathbb{K}$ est un sous-corps de $\mathbb{C}$, "**le rang d'un projecteur est égal à sa trace**".

REMARQUE 4-2.21 De la propriété $\operatorname{trace}(AB) = \operatorname{trace}(BA)$ résulte immédiatement, si A, B et $C \in \mathcal{M}_n(\mathbb{K})$ le fait que $\operatorname{trace}(ABC) = \operatorname{trace}(CAB)$ et plus généralement qu'une permutation circulaire ne change pas la trace d'un produit de matrices carrées. Cependant une permutation quelconque change cette trace en général et il est facile de construire un exemple pour lequel

$$\operatorname{trace}(ABC) \neq \operatorname{trace}(BAC)$$

EXERCICE 4-2.22 Pour $A \in \mathcal{M}_n(\mathbb{K})$, on note φ_A l'application

$$\varphi_A : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K} \qquad M \mapsto \operatorname{trace}\left({}^tAM\right)$$

Montrer que l'application $A \mapsto \varphi_A$ est un isomorphisme entre $\mathcal{M}_n(\mathbb{K})$ et son dual. En déduire qu'une forme linéaire ψ sur $\mathcal{M}_n(\mathbb{K})$ qui vérifie

$$\psi(AB) = \psi(BA)$$

est proportionnelle à la trace.

4-3 Calcul matriciel par blocs

Les notations que nous utilisons dans cette section sont différentes de celles utilisées précédemment. En particulier, le fait de nommer $\mathbb{E}_i$ un espace vectoriel ne signifie plus que cet espace soit de dimension i.

4-3.1 Matrice d'applications linéaires

On se donne deux espaces vectoriels $\mathbb{E}$ et $\mathbb{F}$ décomposés en sommes directes de sous-espaces :

$$\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_j \oplus \cdots \oplus \mathbb{E}_n \quad \text{et} \quad \mathbb{F} = \mathbb{F}_1 \oplus \cdots \oplus \mathbb{F}_i \oplus \cdots \oplus \mathbb{F}_p$$

et on considère un morphisme $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$. On sait déjà (c'est le théorème de recollement linéaire de morphismes) que u est caractérisé par la donnée des morphismes $(u_j)_{1 \leqslant j \leqslant n}$ avec $u_j = u|_{\mathbb{E}_j} \in \mathcal{L}(\mathbb{E}_j, \mathbb{F})$. Dans cette section, nous allons tenir compte également de la décomposition de l'espace d'arrivée. Pour ce faire, on considère la famille de projecteurs $(p_i)_{1 \leqslant i \leqslant p}$ associée à la décomposition

$$\mathbb{F}_1 \oplus \cdots \oplus \mathbb{F}_i \oplus \cdots \oplus \mathbb{F}_p = \mathbb{F}$$

et on considère les morphismes

$$(u_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \quad \text{avec } u_{ij} = p_i \circ u_j = p_i \circ u|_{\mathbb{E}_j}$$

Comme l'image de u_{ij} est incluse dans $\mathbb{F}_i$, on peut, par un abus de langage habituel, considérer que

$$u_{ij} \in \mathcal{L}(\mathbb{E}_j, \mathbb{F}_i)$$

DÉFINITION 4-3.1 *La famille $(u_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$ définie précédemment est appelée matrice d'applications linéaires associée à $u \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ et aux décompositions des espaces $\mathbb{E} = \bigoplus_{j=1}^{n} \mathbb{E}_j$ et $\mathbb{F} = \bigoplus_{i=1}^{p} \mathbb{F}_i$.*

On obtient aisément le raffinement du théorème de recollement :

THÉORÈME 4-3.2 Avec les notations précédentes, la correspondance

$$\mathcal{L}(\mathbb{E}, \mathbb{F}) \to \prod_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \mathcal{L}(\mathbb{E}_j, \mathbb{F}_i) \qquad u \mapsto (u_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$$

est un isomorphisme d'espaces vectoriels.

On reconstitue en effet sans ambiguïté le morphisme u à partir des $(u_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$: si $x \in \mathbb{E}$ est décomposé en $x = x_1 + \cdots + x_j + \cdots + x_n$ avec $x_j \in \mathbb{E}_j$ on a clairement

$$u(x) = \sum_{j=1}^{n} u_j(x_j) = \sum_{j=1}^{n} \sum_{i=1}^{p} u_{ij}(x_j)$$

L'interprétation matricielle de cet isomorphisme est claire, si on travaille dans des bases

$$\mathcal{B} = \bigcup_{j=1}^{n} \mathcal{B}_j \quad \text{et} \quad \mathcal{B}' = \bigcup_{i=1}^{p} \mathcal{B}'_j$$

adaptées aux décompositions de $\mathbb{E}$ et $\mathbb{F}$. Si $A = M(u, \mathcal{B}, \mathcal{B}')$, la matrice A est décomposée "par blocs"

$$A = (A_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$$

où A_{ij} est la matrice extraite de A correspondant aux colonnes de A relatives aux images des vecteurs de $\mathcal{B}_j$ et aux lignes correspondant aux coordonnées selon les vecteurs de $\mathcal{B}'_i$. On a alors clairement

$$A_{ij} = M(u_{ij}, \mathcal{B}_j, \mathcal{B}'_i)$$

4-3.2 Matrice d'applications linéaires associée à un composé de morphismes

Si on intercale entre $\mathbb{E}$ et $\mathbb{F}$ un espace $\mathbb{G}$ lui-même décomposé en une somme directe

$$\mathbb{G} = \bigoplus_{k=1}^{q} \mathbb{G}_k.$$

on s'aperçoit que la composition de morphismes correspond à une règle opératoire sur les matrices d'applications linéaires analogue au produit matriciel classique :

THÉORÈME 4-3.3 Si on se donne deux morphismes u et v

$$\bigoplus_{j=1}^{n} \mathbf{E}_j \xrightarrow{u} \bigoplus_{k=1}^{q} \mathbb{G}_k \xrightarrow{v} \bigoplus_{i=1}^{p} \mathbb{F}_i$$

on a, avec les mêmes notations que précédemment,

$$\forall i \in \{1, \dots, p\} \quad \forall j \in \{1, \dots, n\} \qquad (vu)_{ij} = \sum_{k=1}^{q} v_{ik} u_{kj}$$

Démonstration : Si $x \in \mathbf{E}_j$ on a :

$$vu(x) = v(u_j(x)) = v\left(\sum_{k=1}^{q} u_{kj}(x)\right) = \sum_{k=1}^{q} v(u_{kj}(x))$$

Comme par définition $u_{kj}(x) \in \mathbb{G}_k$ on a aussi

$$vu(x) = \sum_{k=1}^{q} v_k(u_{kj}(x)) = \sum_{k=1}^{q} \sum_{l=1}^{p} v_{lk}(u_{kj}(x))$$

Pour obtenir ensuite $(vu)_{ij}(x)$, il reste à projeter ce vecteur sur $\mathbb{F}_i$ parallèlement à la somme directe des $\mathbb{F}_l$ pour $l \neq i$. On obtient :

$$(vu)_{ij}(x) = \sum_{k=1}^{q} v_{ik}(u_{kj}(x))$$

ce qui démontre la propriété souhaitée. ■

La traduction matricielle est claire : on pourra calculer simplement un produit VU de deux matrices V et U décomposées en blocs rectangulaires si le nombre et la taille des blocs apparaissant en colonnes pour V correspondent au nombre et à la taille des blocs apparaissant en lignes pour U. Formellement tout se passe comme si on manipulait des scalaires au lieu de matrices. Il faut cependant faire attention à ce que, le produit des matrices n'étant pas commutatif, les blocs de V sont à gauche et ceux de U à droite :

$$U = (A_{kj})_{\substack{1\leqslant k\leqslant q\\1\leqslant j\leqslant n}} \text{ et } V = (B_{ik})_{\substack{1\leqslant i\leqslant p\\1\leqslant k\leqslant q}} \Rightarrow VU = (C_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}} \text{ avec } C_{ij} = \sum_{k=1}^{q} B_{ik}A_{kj}$$

4-3.3 Exemple d'utilisation

Les deux exercices qui suivent peuvent se résoudre en utilisant des calculs par blocs :

EXERCICE 4-3.4 Soient A et B deux matrices de $\mathcal{GL}_n(\mathbb{K})$ qui commutent. Montrer que

$$M = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathcal{GL}_{2n}(\mathbb{K}) \Leftrightarrow A^2 + B^2 \in \mathcal{GL}_n(\mathbb{K})$$

On pourra chercher les vecteurs Z de $\mathbb{K}^{2n}$ qui sont dans le noyau de M en les décomposant, comme le suggère la forme de M en deux blocs de taille n : $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$.

EXERCICE 4-3.5 Soit $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p)$ de rang r. Déterminer le rang de l'application linéaire

$$\mathcal{L}_{\mathbb{K}}(\mathbb{F}_p, \mathbb{E}_n) \overset{\Psi}{\to} \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n, \mathbb{F}_p) \qquad v \mapsto u \circ v \circ u$$

L'idée est probablement de déterminer le noyau de cette application. On peut travailler matriciellement, en choisissant un couple de bases où la donnée u s'exprime le plus simplement possible. On peut aussi raisonner géométriquement, en remarquant que

$$v \in \ker \Psi \Leftrightarrow v(\operatorname{Im} u) \subset \ker u$$

et en montrant que ce noyau est isomorphe à $\mathcal{L}(\operatorname{Im} u, \ker u) \times \mathcal{L}(\mathbb{S}, \mathbb{E}_n)$ où $\mathbb{S}$ est un supplémentaire de $\operatorname{Im} u$ dans $\mathbb{F}_p$.

4-4 Déterminants

Le déterminant est un outil **théorique** important pour étudier l'inversibilité d'une matrice carrée : il montre que les matrices non inversibles de $\mathcal{M}_n(\mathbb{K})$ sont les solutions de l'équation **polynomiale** (par rapport aux n^2 "variables" a_{ij} coefficients de la matrice A)

$$\det A = 0$$

L'existence de cette condition polynomiale a de multiples conséquences. L'une d'elles est notamment, si on travaille sur $\mathbb{R}$, qu'une légère perturbation sur une matrice inversible donne encore une matrice inversible. Nous y reviendrons dans le chapitre de topologie des espaces normés. Si on identifie $\mathcal{M}_n(\mathbb{R})$ et $\mathbb{R}^{n^2}$, on peut aussi interpréter l'ensemble des matrices non inversibles comme un (hyper)surface ayant une équation polynomiale (on parle de surface algébrique). On se rend alors compte que $\mathcal{M}_n(\mathbb{R})$ n'est (presque) constitué que de matrices inversibles.

Dans le cadre particulier des espaces euclidiens de dimension n orientés, nous interpréterons ultérieurement le produit mixte de n vecteurs (c'est-à-dire leur déterminant dans une base orthonormale directe) comme un volume (algébrique), celui du parallélépipède construit sur ces vecteurs.

4-4.1 Applications multilinéaires symétriques et antisymétriques

Dans cette section, tous les espaces vectoriels sont sur le même corps commutatif $\mathbb{K}$.

4-4.1.1 Applications multilinéaires

DÉFINITION 4-4.1 *Si $(\mathbb{E}_i)_{1\leqslant i\leqslant p}$ sont p espaces vectoriels, et $\mathbb{F}$ est lui aussi un espace vectoriel, une application*

$$\Phi : \mathbb{E}_1 \times \cdots \times \mathbb{E}_p \to \mathbb{F}$$

est dite p-linéaire ssi elle est linéaire par rapport à chacun de ses arguments, ce qui signifie que les applications partielles

$$\mathbb{E}_i \to \mathbb{F} \qquad x_i \mapsto \Phi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

obtenues en fixant $p-1$ des arguments de Φ sont linéaires.

Nous avons déjà rencontré le crochet de dualité (section **3-1.1**) et le produit matriciel (*cf.* théorème **4-1**.12) comme exemples d'applications bilinéaires. Nous rencontrerons de multiples exemples ultérieurement. Disons simplement que la notion de multilinéarité est une généralisation de la notion de "produit", l'application p linéaire la plus simple à envisager étant l'application

$$\mathbb{K}^p \to \mathbb{K} \qquad (x_1, \ldots, x_p) \mapsto x_1 \cdots x_p$$

Il ne faut évidemment pas confondre linéarité et p-linéarité : si

$$\Phi : \mathbb{E}_1 \times \cdots \times \mathbb{E}_p \to \mathbb{F}$$

est linéaire, le développement de $\Phi(x_1 + y_1, \ldots, x_p + y_p)$ donne une somme de deux termes. Si Φ est p-linéaire, ce développement donne 2^p termes. De même la p linéarité entraîne

$$\Phi(\lambda x_1, \ldots, \lambda x_p) = \lambda^p \Phi(x_1, \ldots, x_p) \text{ et } \Phi(x_1, \ldots, 0_{\mathbb{E}_i}, \ldots, x_p) = 0_{\mathbb{F}}$$

THÉORÈME 4-4.2 L'ensemble des applications p linéaires de $\mathbb{E}_1 \times \cdots \times \mathbb{E}_p$ dans $\mathbb{F}$ est un espace vectoriel qu'on note $\mathcal{L}_p(\mathbb{E}_1 \times \cdots \times \mathbb{E}_p, \mathbb{F})$.

C'est en effet clairement un sous-espace vectoriel de $\mathbb{F}^{\mathbb{E}_1 \times \cdots \times \mathbb{E}_p}$. Lorsque tous les $\mathbb{E}_i$ sont égaux à un même espace $\mathbb{E}$, on note (s'il n'y a pas d'ambiguïté) $\mathcal{L}_p(\mathbb{E}, \mathbb{F})$ pour $\mathcal{L}_p(\mathbb{E}^p, \mathbb{F})$, et on parle d'application p-linéaire définie sur $\mathbb{E}$ (bien que le domaine de définition soit $\mathbb{E}^p$).

4-4.1.2 Symétrie et antisymétrie

On se place ici dans la situation où les $\mathbb{E}_i$ sont tous égaux à un même espace $\mathbb{E}$ et on suppose que $\mathbb{F} \neq \{0_{\mathbb{F}}\}$.

DÉFINITION 4-4.3 *Une application $\Phi \in \mathcal{L}_p(\mathbb{E}, \mathbb{F})$ est dite*

- *symétrique si et seulement si la valeur de $\Phi(x_1, \ldots, x_p)$ ne change pas si on transpose deux vecteurs dans le p-uplet $(x_1, \ldots, x_p)$.*
- *antisymétrique si et seulement si la valeur de $\Phi(x_1, \ldots, x_p)$ est changée en son opposé si on transpose deux vecteurs dans le p-uplet $(x_1, \ldots, x_p)$*

On note $\mathcal{L}_{p,s}(\mathbb{E}, \mathbb{F})$ (resp. $\mathcal{L}_{p,a}(\mathbb{E}, \mathbb{F})$) l'ensemble des applications p-linéaires symétriques (resp. antisymétriques) de $\mathbb{E}$ dans $\mathbb{F}$. Ce sont clairement deux sous-espaces vectoriels de $\mathcal{L}_p(\mathbb{E}, \mathbb{F})$.

Dans le cas très particulier où le corps $\mathbb{K}$ est de caractéristique 2 (c'est-à-dire $2.1_{\mathbb{K}} = 0_{\mathbb{K}}$, soit $1_{\mathbb{K}} = -1_{\mathbb{K}}$) ces deux sous-espaces sont confondus. Il n'y a dans ce cas que la notion de symétrie qui est intéressante. Dans les autres cas (donc en particulier lorsque $\mathbb{K}$ est un sous-corps de $\mathbb{C}$), les sous-espaces $\mathcal{L}_{p,s}(\mathbb{E}, \mathbb{F})$ et $\mathcal{L}_{p,a}(\mathbb{E}, \mathbb{F})$ sont deux sous-espaces indépendants de $\mathcal{L}_p(\mathbb{E}, \mathbb{F})$.

Rappel : On suppose que p est un entier naturel $\geqslant 2$. On note $(S_p, \circ)$ le groupe des permutations de l'ensemble $\{1, \ldots, p\}$. Il est de cardinal $p!$. Une permutation $\tau \in S_p$ est appelée transposition si elle permute deux éléments distincts i et j (on note $\tau = (i, j)$ cette transposition) et laisse invariants tous les autres . On démontre que toute permutation $\sigma \in S_p$ peut s'écrire comme produit de transpositions et que la parité du nombre de transpositions apparaissant dans une telle décomposition de σ ne dépend pas de la décomposition. On dit alors que σ est une permutation paire si elle est produit d'un nombre pair de transpositions et on pose $\varepsilon(\sigma) = 1$. Si, au contraire, σ est produit d'un nombre impair de transpositions, σ est dite impaire et on pose $\varepsilon(\sigma) = -1$.

$$\varepsilon : (S_p, \circ) \to (\{-1, +1\}, \times)$$

est un morphisme de groupes qu'on appelle signature[1].

La définition d'application symétrique ou antisymétrique a été donnée en faisant opérer une transposition sur la famille de vecteurs $(x_i)_{1\leqslant i\leqslant p}$. Il en résulte immédiatement

PROPOSITION 4-4.4 Si $\Phi \in \mathcal{L}_p(\mathbb{E},\mathbb{F})$, $\sigma \in S_p$ et $(x_i)_{1\leqslant i\leqslant p}$ sont p vecteurs de $\mathbb{E}$:

$$\Phi \text{ symétrique } \Rightarrow \Phi\left(x_{\sigma(1)},\ldots,x_{\sigma(p)}\right) = \Phi(x_1,\ldots,x_p)$$

$$\Phi \text{ antisymétrique } \Rightarrow \Phi\left(x_{\sigma(1)},\ldots,x_{\sigma(p)}\right) = \varepsilon(\sigma)\,\Phi(x_1,\ldots,x_p)$$

4-4.1.3 Application multilinéaire alternée

Nous donnons ici une définition qui s'avèrera être équivalente à la notion d'antisymétrie **lorsque le corps de base est de caractéristique différente de 2.**

DÉFINITION 4-4.5 *Une application $\Phi \in \mathcal{L}_p(\mathbb{E},\mathbb{F})$ est dite alternée ssi on a $\Phi(x_1,\ldots,x_p) = 0_{\mathbb{F}}$ dès que deux (au moins) des vecteurs de la famille $(x_i)_{1\leqslant i\leqslant p}$ sont égaux.*

PROPOSITION 4-4.6 Si $\Phi \in \mathcal{L}_p(\mathbb{E},\mathbb{F})$ est alternée, elle est antisymétrique. La réciproque est vraie lorsque le corps de base est de caractéristique différente de 2.

Démonstration : Si $\Phi \in \mathcal{L}_p(\mathbb{E},\mathbb{F})$ est alternée, et si $(x_i)_{1\leqslant i\leqslant p} \in \mathbb{E}$, en développant par multilinéarité

$$\Phi\left(x_1,\ldots,\underset{i^{\text{ème}}\text{place}}{\underset{\downarrow}{x_i+x_j}},\ldots,\underset{j^{\text{ème}}\text{place}}{\underset{\downarrow}{x_i+x_j}},\ldots,x_p\right)$$

(qui est nul par hypothèse), on obtient une somme de quatre termes dont deux sont nuls et il reste

$$\Phi\left(x_1,\ldots,\underset{\text{place } i}{\underset{\downarrow}{x_i}},\ldots,\underset{\text{place } j}{\underset{\downarrow}{x_j}},\ldots,x_p\right) = -\Phi\left(x_1,\ldots,\underset{\text{place } i}{\underset{\downarrow}{x_j}},\ldots,\underset{\text{place } j}{\underset{\downarrow}{x_i}},\ldots,x_p\right)$$

ce qui montre bien que Φ est antisymétrique.

[1]On montre aussi que

$$\varepsilon(\sigma) = \prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}$$

et vaut -1 ssi σ présente un nombre impair d'inversions, c'est à dire de couples (i,j) tels que $i<j$ et $\sigma(i)>\sigma(j)$.

Réciproquement, si Φ est antisymétrique et si $(x_k)_{1\leqslant k\leqslant p}$ est une famille de vecteurs de $\mathbb{E}$ avec $x_i = x_j$ pour $i \neq j$, en faisant opérer sur cette famille la transposition $\tau = (i, j)$, on obtient

$$\Phi(x_1, \dots, x_p) = -\Phi(x_1, \dots, x_p)$$

ce qui donne bien $\Phi(x_1, \dots, x_p) = 0_{\mathbb{F}}$, lorsque la caractéristique du corps est différente de 2. ■

Les propriétés qui suivent sont à la base de la théorie des déterminants :

PROPOSITION 4-4.7 Si $\Phi \in \mathcal{L}_p(\mathbb{E}, \mathbb{F})$ est alternée, on ne change pas la valeur de $\Phi(x_1, \dots, x_p)$ en ajoutant à un des vecteurs une combinaison linéaire des autres. Si $(x_1, \dots, x_p)$ est un système lié, alors $\Phi(x_1, \dots, x_p) = 0_{\mathbb{F}}$.

Démonstration : Par exemple, on a

$$\Phi\left(x_1 + \sum_{i=2}^{p} \lambda_i x_i, \dots, x_p\right) =$$

$$\Phi(x_1, \dots, x_p) + \sum_{i=2}^{p} \lambda_i \Phi(x_i, \dots, x_i, \dots, x_p) = \Phi(x_1, \dots, x_p)$$

En particulier, si le système est lié, on peut faire apparaître le vecteur nul en ajoutant à un des vecteurs (bien choisi) une combinaison linéaire des autres. On en déduit alors $\Phi(x_1, \dots, x_p) = 0_{\mathbb{F}}$. ■

4-4.1.4 Exercice : symétrisée et antisymétrisée d'une application p-linéaire

A partir d'une application $\Phi \in \mathcal{L}_p(\mathbb{E}, \mathbb{F})$, il est facile de construire théoriquement des applications symétriques ou antisymétriques. Il faut cependant prendre garde que le résultat obtenu peut parfois être l'application identiquement nulle, ce qui n'a alors aucun intérêt.

DÉFINITION 4-4.8 *Si $\Phi \in \mathcal{L}_p(\mathbb{E}, \mathbb{F})$, on appelle symétrisée de Φ l'application*

$$\Phi^s : \mathbb{E}^p \to \mathbb{F} \qquad \Phi^s(x_1, \dots, x_p) = \sum_{\sigma \in S_p} \Phi\left(x_{\sigma(1)}, \dots, x_{\sigma(p)}\right)$$

et antisymétrisée de Φ l'application

$$\Phi^a : \mathbb{E}^p \to \mathbb{F} \qquad \Phi^a(x_1, \dots, x_p) = \sum_{\sigma \in S_p} \varepsilon(\sigma) \Phi\left(x_{\sigma(1)}, \dots, x_{\sigma(p)}\right)$$

Vérifier que ces applications sont p-linéaires, que Φ^s est symétrique et Φ^a est antisymétrique.

La seule difficulté est sans doute de montrer que Φ^a est antisymétrique. Si τ est une transposition, on a

$$\Phi^a\left(x_{\tau(1)}, \dots, x_{\tau(p)}\right) = \sum_{\sigma \in S_p} \varepsilon(\sigma) \Phi\left(x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(p))}\right)$$

et on utilise ensuite le fait que $\sigma \mapsto \tau \circ \sigma$ est une bijection de S_p dans lui-même, en remarquant également que $\varepsilon(\tau \circ \sigma) = \varepsilon(\tau)\varepsilon(\sigma) = -\varepsilon(\sigma)$.

REMARQUE 4-4.9 Si Φ est antisymétrique, on a évidemment $\Phi^a = p!\,\Phi$. De même si Φ est symétrique $\Phi^s = p!\,\Phi$. (Attention cependant au terme $p!$ qui peut être nul en caractéristique différente de 0)

REMARQUE 4-4.10 Si $\mathbb{K}$ est de caractéristique 2, il est facile de voir que Φ^a est en fait alternée, car si deux vecteurs sont égaux, chaque terme de la somme apparaît en fait deux fois dans celle-ci et, en regroupant les termes 2 par 2, le résultat est clairement nul. Ceci entraîne que la théorie des déterminants développée ultérieurement est en fait aussi valable en caractéristique 2.

4-4.2 Définition et propriétés des déterminants

4-4.2.1 Déterminant d'un système de n vecteurs dans une base de $\mathbb{E}_n$, d'une matrice carrée

Soit $\mathbb{E}_n$ un espace de dimension n rapporté à une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$. Une famille $\mathcal{F} = (x_i)_{1 \leqslant i \leqslant n}$ de n vecteurs de $\mathbb{E}_n$ est caractérisée par sa matrice

$$M_{\mathcal{B}}(\mathcal{F}) = A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(\mathbb{K})$$

L'application

$$\mathbb{E}^n \xrightarrow{\Phi} \mathbb{K} \qquad (x_1, \ldots, x_n) \mapsto a_{11}a_{22} \cdots a_{nn}$$

est n linéaire, à valeurs dans $\mathbb{K}$. On parle alors de **forme n-linéaire**. Le déterminant dans la base $\mathcal{B}$ est, par définition, l'antisymétrisée de cette application :

THÉORÈME 4-4.11 (et définition) Avec les notations précédentes, on appelle déterminant dans la base $\mathcal{B}$ l'application

$$\det_{\mathcal{B}} : \mathbb{E}_n^n \to \mathbb{K} \quad (x_1, \ldots, x_n) \mapsto \sum_{\sigma \in S_n} \varepsilon(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

C'est une forme n-linéaire alternée sur l'espace $\mathbb{E}_n$ de dimension n. Elle n'est pas identiquement nulle car

$$\det_{\mathcal{B}}(\mathcal{B}) = 1$$

Démonstration : Seule la propriété $\det_{\mathcal{B}}(\mathcal{B}) = 1$ est à voir. Lorsque $\mathcal{F} = \mathcal{B}$, la matrice A est égale à I_n, et pour qu'un produit

$$a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

soit non nul, il est clair qu'il est nécessaire d'avoir $\sigma = id_{\{1,\ldots,n\}}$, auquel cas ce produit vaut 1. ■

Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(\mathbb{K})$, l'expression $\sum_{\sigma \in S_n} \varepsilon(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ est donc le déterminant des vecteurs colonnes de A dans la base canonique de $\mathbb{K}^n$. Ce déterminant est aussi appelé déterminant de la matrice carrée A et est noté $\det A$.

DÉFINITION 4-4.12 *Le déterminant d'une matrice carrée est le déterminant de ses vecteurs colonnes dans la base canonique de* $\mathbb{K}^n$. *On note*

$$\det A = \det (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} = |a_{ij}|_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$$

Cette définition aurait pu être donnée sur les lignes, comme le montre la règle de dualité :

THÉORÈME 4-4.13 Si $A \in \mathcal{M}_n(\mathbb{K})$

$$\det A = \det {}^tA$$

Démonstration : Posons $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$. Si $\sigma \in S_n$, on a

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}$$

à cause de la commutativité de la multiplication dans $\mathbb{K}$: comme σ est bijective, dans le produit $a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$, il y a pour tout $i \in \{1,\dots,n\}$ un seul terme dont le second indice vaut i, c'est $a_{\sigma^{-1}(i)i}$. Comme $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$, on a

$$\det A = \sum_{\sigma\in S_n} \varepsilon\left(\sigma^{-1}\right) a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}$$

et on conclut aisément en remarquant que $\sigma \to \sigma^{-1}$ est bijective de S_n dans lui-même. ■

Pourquoi cette expression tarabiscotée et parachutée du déterminant ? Parce qu'elle apparaît de manière naturelle dans le calcul mené dans la section qui suit :

4-4.2.2 Espace $\mathcal{A}_n(\mathbb{E}_n)$: résultat fondamental

On note $\mathcal{A}_n(\mathbb{E}_n)$ l'espace $\mathcal{L}_n(\mathbb{E}_n, \mathbb{K})$ des formes n-linéaires alternées sur l'espace $\mathbb{E}_n$ de dimension n. Les démonstrations sur les déterminants s'inspirent très souvent du résultat suivant :

THÉORÈME 4-4.14 L'espace $\mathcal{A}_n(\mathbb{E}_n)$ est un espace vectoriel de dimension 1. Plus précisément, si $\mathcal{B}$ est une base arbitraire de $\mathbb{E}_n$, toute forme n-linéaire alternée sur l'espace $\mathbb{E}_n$ est proportionnelle à $\det_{\mathcal{B}}$. Si $\Phi \in \mathcal{A}_n(\mathbb{E}_n)$

$$\Phi = \Phi(\mathcal{B}) \det{}_{\mathcal{B}}$$

Démonstration : On sait déjà que $\mathcal{A}_n(\mathbb{E}_n)$ est un espace vectoriel non réduit à la forme nulle, puisqu'il contient le déterminant dans une base arbitraire de $\mathbb{E}_n$. Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une telle base et soient $\Phi \in \mathcal{A}_n(\mathbb{E}_n)$ et $(x_i)_{1\leqslant i\leqslant n}$ une famille de n vecteurs de $\mathbb{E}_n$ caractérisés par

$$M_{\mathcal{B}}(x_1,\dots,x_n) = A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} \in \mathcal{M}_n(\mathbb{K})$$

On a alors

$$\Phi(x_1,\ldots,x_n) = \Phi\left(\sum_{i_1=1}^{n} a_{i_1,1}e_{i_1},\ldots,\sum_{i_n=1}^{n} a_{i_n,n}e_{i_n}\right)$$

En développant par n-linéarité, on obtient une somme de n^n termes :

$$\Phi(x_1,\ldots,x_n) = \sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n} a_{i_1,1}\cdots a_{i_n,n}\Phi(e_{i_1},\ldots,e_{i_n})$$

Mais comme Φ est alternée, il ne subsiste que $n!$ termes éventuellement non nuls : ceux pour lesquels les indices $i_1,\ldots,i_n$ sont distincts deux à deux, c'est-à-dire pour lesquels il existe $\sigma \in S_n$ avec $\forall k \quad i_k = \sigma(k)$. En utilisant σ comme indice, on obtient :

$$\Phi(x_1,\ldots,x_n) = \sum_{\sigma\in S_n} a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}\Phi\left(e_{\sigma(1)},\ldots,e_{\sigma(n)}\right)$$

soit enfin, puisque Φ est antisymétrique

$$\Phi(x_1,\ldots,x_n) = \Phi(e_1,\ldots,e_n)\det_{\mathcal{B}}(x_1,\ldots,x_n)$$

ce qui démontre le résultat. ■

EXERCICE 4-4.15 Montrer que, si $p \leqslant n$, $\mathcal{A}_p(\mathbb{E}_n) = \mathcal{L}_p(\mathbb{E}_n,\mathbb{K})$ est un espace de dimension C_n^p. Que se passe-t-il si $p > n$?

Le théorème précédent a deux conséquences immédiates :

COROLLAIRE 4-4.16 Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $\mathbb{E}_n$

$$\det_{\mathcal{B}'} = \det_{\mathcal{B}'}(\mathcal{B})\det_{\mathcal{B}}$$

COROLLAIRE 4-4.17 Une famille de n vecteurs de $\mathbb{E}_n$ est libre si et seulement si son déterminant dans une base arbitraire de $\mathbb{E}_n$ est non nul.

Démonstration : Soit $\mathcal{F}$ une famille de cardinal n de $\mathbb{E}_n$ et $\mathcal{B}$ une base. Si $\mathcal{F}$ est liée, on a évidemment $\det_{\mathcal{B}}(\mathcal{F}) = 0$ (*cf.* proposition **4-4**.7). Si $\mathcal{F}$ est libre de cardinal n, c'est une base de $\mathbb{E}_n$ et donc

$$\det_{\mathcal{F}}(\mathcal{F}) = \det_{\mathcal{F}}(\mathcal{B})\det_{\mathcal{B}}(\mathcal{F}) = 1$$

ce qui donne évidemment $\det_{\mathcal{B}}(\mathcal{F}) \neq 0$. ■

Comme une matrice carrée est inversible si et seulement si ses vecteurs colonnes sont indépendants, on a aussi

COROLLAIRE 4-4.18 Une matrice carrée est inversible si et seulement si son déterminant est non nul.

Attention ! Le résultat précédent est essentiellement théorique. Dans la pratique, on utilise rarement le déterminant pour montrer qu'une matrice est inversible, à cause de la difficulté engendrée par le calcul des déterminants.

4-4.2.3 Développement suivant une ligne ou une colonne

Il faut avant tout se souvenir que, par construction, le déterminant est une forme n-linéaire alternée sur l'espace des vecteurs colonnes (ou lignes). Donc

> **On ne change pas la valeur d'un déterminant en ajoutant à une colonne (ligne) une combinaison linéaire des autres colonnes (lignes). Le déterminant est changé en son opposé si on permute deux lignes ou colonnes. Le déterminant est n-linéaire par rapport à ses colonnes (lignes).**

L'utilisation de ces propriétés va permettre un développement théorique du déterminant :

Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} = (C_1, \dots, C_n)$ une matrice de $\mathcal{M}_n(\mathbb{K})$. On note $(E_i)_{1\leqslant i\leqslant n}$ la base canonique de $\mathbb{K}^n$ (écrite sous forme de vecteurs colonnes). On a alors pour tout $j \in \{1, \dots, n\}$:

$$\det A = \det(C_1, \dots, C_j, \dots, C_n) = \det\left(C_1, \dots, \sum_{i=1}^{n} a_{ij}E_i, \dots, C_n\right)$$

Par linéarité du déterminant par rapport à sa $j^{\text{ème}}$ colonne, on a

$$\det A = \sum_{i=1}^{n} a_{ij} \det(C_1, \dots, E_i, \dots, C_n) = \sum_{i=1}^{n} a_{ij}C_{ij}$$

où C_{ij} est le **cofacteur de l'élément** (i, j) (ou cofacteur d'indices (i, j)) de la matrice A, c'est-à-dire le déterminant de la matrice X_{ij} obtenue en remplaçant la $j^{\text{ème}}$ colonne de A par la $i^{\text{ème}}$ colonne de la base canonique. Par $n-i$ transpositions successives de lignes, on peut faire descendre la $i^{\text{ème}}$ ligne de X_{ij} en dernière position, puis par $n-j$ transpositions de colonnes amener la $j^{\text{ème}}$ colonne de la matrice ainsi obtenue en dernière place. On obtient donc

$$C_{ij} = (-1)^{n-i}(-1)^{n-j}\begin{vmatrix} & & 0 \\ & A_{ij} & \vdots \\ & & 0 \\ a_{i1} & \cdots\ a_{in} & 1 \end{vmatrix} = (-1)^{i+j}\,\Delta_{ij}$$

où A_{ij} est la matrice de $\mathcal{M}_{n-1}(\mathbb{K})$ extraite de A en rayant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne, et où $a_{i1}\cdots a_{in}$ représente la $i^{\text{ème}}$ ligne de A dont on a enlevé le coefficient a_{ij} d'indice j.

Si $B = (b_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}$ est une matrice décomposée par blocs

$$B = \begin{pmatrix} & & 0 \\ & B' & \vdots \\ & & 0 \\ b_{n1} & \cdots\ b_{nn-1} & 1 \end{pmatrix}$$

dans la formule

$$\det B = \sum_{\sigma \in S_n} \varepsilon(\sigma)\, b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n}$$

les seuls termes éventuellement non nuls sont ceux pour lesquels $\sigma(n) = n$. On peut donc considérer que σ', restriction de σ à $\{1, \ldots, n-1\}$, est dans S_{n-1}, et comme une décomposition de σ' en produit de transpositions donne aussi une décomposition de σ, on a finalement

$$\det B = \sum_{\sigma' \in S_{n-1}} \varepsilon(\sigma')\, b_{\sigma'(1)1} b_{\sigma'(2)2} \cdots b_{\sigma'(n-1)n-1} = \det B'$$

On obtient donc :

THÉORÈME 4-4.19 Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(\mathbb{K})$, pour tout $j \in \{1, \ldots, n\}$ on a l'égalité

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij}$$

(développement suivant la $j^{\text{ème}}$ colonne). Dans cette formule, le cofacteur C_{ij} est égal à

$$C_{ij} = (-1)^{i+j} \Delta_{ij}$$

où Δ_{ij} est le déterminant (d'ordre $n-1$) de la matrice extraite de A en rayant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne. Ce déterminant Δ_{ij} est appelé (déterminant) mineur d'indices (i, j) de la matrice A. Par dualité, on a évidemment, pour tout $i \in \{1, \ldots, n\}$

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$$

(développement suivant la $i^{\text{ème}}$ ligne).

Telle quelle, cette formule est essentiellement théorique, puisqu'elle ramène le calcul d'un déterminant d'ordre n à celui de n déterminants d'ordre $n-1$, donc $n(n-1)$ déterminants d'ordre $n-2$ ... et finalement à $n!$ déterminants d'ordre 1, ce qui correspond exactement à la formule de définition, $n!$ étant le cardinal du groupe S_n. Le calcul sur ordinateur d'un déterminant d'ordre 30 par cette formule amènerait plus de 10^{30} multiplications ($n! > \left(\frac{n}{e}\right)^n$) ce qui, compte tenu du temps requis pour une opération élémentaire donnerait un temps de calcul supérieur à 10^{12} ans (évaluation grossière).

Lorsqu'on utilise cette formule, on essaie d'abord, en utilisant les propriétés du déterminant, de faire apparaître un maximum de zéros sur une ligne ou une colonne, par des opérations de combinaisons linéaires, ce qui diminue le nombre de cofacteurs à calculer effectivement.

EXERCICE 4-4.20 Calculer le déterminant de la matrice

$$M_n = \begin{pmatrix} \alpha & 0 & \cdots & \cdots & 0 & \beta \\ 0 & \searrow & 0 & 0 & \nearrow & 0 \\ \vdots & 0 & \alpha & \beta & 0 & \vdots \\ \vdots & 0 & \beta & \alpha & 0 & \vdots \\ 0 & \nearrow & 0 & 0 & \searrow & 0 \\ \beta & 0 & \cdots & \cdots & 0 & \alpha \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{K})$$

4-4.2.4 Expression théorique d'une matrice inverse

DÉFINITION 4-4.21 *Si $A \in \mathcal{M}_n(\mathbb{K})$, on appelle comatrice de A et on note* $\operatorname{com} A$ *la matrice*

$$\operatorname{com} A = (C_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

des cofacteurs de la matrice A.

THÉORÈME 4-4.22 Si $A \in \mathcal{M}_n(\mathbb{K})$, on a

$$A\, {}^t\!\operatorname{com} A = {}^t\!\operatorname{com} A.A = (\det A)\, I_n$$

Démonstration : Le coefficient de la $i^{\text{ème}}$ ligne et $i^{\text{ème}}$ colonne de la matrice $B = A\, {}^t\!\operatorname{com} A$ vaut

$$b_{ii} = \sum_{k=1}^{n} a_{ik} C_{ik}$$

et est donc égal à $\det A$ (développement de $\det A$ par rapport à la $i^{\text{ème}}$ ligne). Par contre, si $i \neq j$,

$$b_{ij} = \sum_{k=1}^{n} a_{ik} C_{jk}$$

et ce terme est nul, puisqu'il correspond au développement par rapport à la $j^{\text{ème}}$ ligne du déterminant de la matrice obtenue à partir de A en remplaçant la $j^{\text{ème}}$ ligne par la $i^{\text{ème}}$. Cette matrice ayant deux lignes égales a un déterminant nul. On calculerait de la même façon le produit ${}^t\!\operatorname{com} A.A$. ■

Comme on sait déjà qu'une matrice carrée est inversible si et seulement si son déterminant est non nul, on obtient :

COROLLAIRE 4-4.23 Si $A \in \mathcal{GL}_n(\mathbb{K})$ (avec $n \geqslant 2$), on a

$$A^{-1} = \frac{1}{\det A}\, {}^t\!\operatorname{com} A$$

Dans le cas $n = 2$, on obtient la formule (qu'il est intéressant de retenir)

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

lorsque, bien entendu, $ad - bc \neq 0$. Hormis ce cas particulier, il est extrêmement rare que l'on utilise cette formule pour calculer effectivement l'inverse d'une matrice. Le résultat qui précède est essentiellement théorique, et est le plus souvent traduit par le corollaire suivant, qui affirme l'existence d'une formule d'un type particulier pour expliciter A^{-1}, alors que la formule en elle-même n'est pas fondamentale.

COROLLAIRE 4-4.24 Sur $\mathcal{GL}_n(\mathbb{K})$, les coefficients de la matrice M^{-1} sont des fonctions rationnelles (quotients de deux polynômes à n^2 variables) des coefficients de la matrice M.

En effet, tous les cofacteurs ainsi que le déterminant de M s'expriment comme fonctions polynomiales des coefficients de M. En particulier, lorsque $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$, cela entraînera que la matrice M^{-1} dépend "continûment" de M.

EXERCICE 4-4.25 Montrer que

$$\det(\operatorname{com} A) = (\det A)^{n-1}$$

EXERCICE 4-4.26 Discuter, en fonction du rang de la matrice A, le rang de la comatrice de A. (Indication : quel théorème peut-on utiliser pour déterminer le rang d'un morphisme ?)

REMARQUE 4-4.27 La détermination pratique de l'inverse d'une matrice peut se faire en utilisant la méthode du pivot de Gauss. Cette méthode, efficace lorsque les coefficients de la matrice sont "numériques", c'est-à-dire ne comportent pas de "paramètres", a cependant un gros défaut : elle rompt parfois une "symétrie" dans la matrice, puisqu'elle fonctionne en faisant jouer un rôle à une ligne particulière (celle dont l'élément diagonal est choisi comme pivot). On peut parfois raisonner en respectant la symétrie. Quelle que soit la méthode suivie, elle correspond finalement à la recherche de l'antécédent d'un vecteur Y quelconque de $\mathbb{K}^n$ par le morphisme canoniquement associé à la matrice A, c'est-à-dire à la résolution d'un système :

$$AX = Y \Leftrightarrow X = A^{-1}Y$$

EXERCICE 4-4.28 Si les a_i sont des scalaires, inverser la matrice (lorsque c'est possible)

$$\begin{pmatrix} 1+a_1 & 1 & \cdots & \cdots & 1 \\ 1 & \ddots & 1 & & \vdots \\ \vdots & 1 & 1+a_i & 1 & \vdots \\ \vdots & & 1 & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & 1+a_n \end{pmatrix}$$

4-4.3 Déterminant d'un endomorphisme

Nous commencerons par donner une définition géométrique du déterminant, avant de le relier à la matrice dans une base quelconque.

THÉORÈME 4-4.29 (et définition) Soit $\mathbb{E}_n$ un espace de dimension n et $u \in \mathcal{L}(\mathbb{E}_n)$. Il existe un unique scalaire, qu'on appelle déterminant de u et que l'on note $\det u$ tel que pour tout base $\mathcal{B}$ de $\mathbb{E}_n$ et tout système $(x_i)_{1 \leqslant i \leqslant n}$ de n vecteurs de $\mathbb{E}_n$ on ait

$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = \det u . \det_{\mathcal{B}}(x_1, \dots, x_n)$$

Démonstration : Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $\mathbb{E}_n$, les applications $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ sont proportionnelles. Il suffit clairement de faire la vérification pour une base particulière. On remarque que l'application

$$(x_1, \dots, x_n) \mapsto \det_{\mathcal{B}}(u(x_1), \dots, u(x_n))$$

est dans $\mathcal{A}_n(\mathbb{E}_n)$ (vérification immédiate). Comme cet espace est de dimension 1 et est engendré par $\det_{\mathcal{B}}$, le résultat en découle. ■

REMARQUE 4-4.30 Une autre manière d'énoncer le résultat précédent est de dire que l'application

$$\mathcal{A}_n(\mathbb{E}_n) \to \mathcal{A}_n(\mathbb{E}_n) \quad \Phi \mapsto \Phi \circ u$$

(où $\Phi \circ u$ représente, avec un petit abus de notation, l'application définie par $(x_1, \dots, x_n) \mapsto \Phi(u(x_1), \dots, u(x_n))$) est un endomorphisme d'un espace de dimension 1 et ne peut donc qu'être une homothétie, dont le rapport est par définition le déterminant de u.

REMARQUE 4-4.31 Avec l'interprétation géométrique que l'on donnera dans le cadre de la théorie des espaces euclidiens, le déterminant d'un endomorphisme apparaîtra comme le rapport de deux "volumes" : celui du parallélépipède construit sur les vecteurs images $(u(x_1), \dots, u(x_n))$ et celui du parallélépipède construit sur les vecteurs $(x_1, \dots, x_n)$. Si on note $\mathcal{P}$ ce dernier, on aura donc

$$\det u = \frac{\mathrm{vol}\, u(\mathcal{P})}{\mathrm{vol}\, \mathcal{P}}$$

C'est pour cela qu'il ne sera pas étonnant de retrouver des déterminants dans les formules de changement de variables pour les intégrales multiples.

En particulier, en prenant pour vecteurs $(x_i)_{1 \leqslant i \leqslant n}$ les vecteurs de la base $\mathcal{B}$, on obtient

$$\det u = \det_{\mathcal{B}}(u(\mathcal{B})) \text{ (indépendant de la base } \mathcal{B})$$

Comme les colonnes de la matrice $M(u, \mathcal{B})$ donnent exactement les coordonnées des vecteurs de $u(\mathcal{B})$ dans $\mathcal{B}$, on a donc

$$\det u = \det M(u, \mathcal{B})$$

COROLLAIRE 4-4.32 Sur l'espace $\mathcal{M}_n(\mathbb{K})$, le déterminant est un invariant de similitude.

COROLLAIRE 4-4.33 Un endomorphisme de $\mathbb{E}_n$ est un automorphisme si et seulement si son déterminant est non nul.

Pour dire les choses autrement, si $u \in \mathcal{L}(\mathbb{E}_n)$ est non inversible, son image est de dimension inférieure ou égale à $n-1$. L'endomorphisme u "aplatit" donc tous les parallélépipèdes construits sur n vecteurs. Ce n'est pas le cas si u est un automorphisme.

La remarque **4-4**.30 montre que la composition par u est en fait une homothétie de $\mathcal{A}_n(\mathbb{E}_n)$ de rapport $\det u$. Comme composer les homothéties, c'est multiplier les rapports, on a :

THÉORÈME 4-4.34 Si u et $v \in \mathcal{L}(\mathbb{E}_n)$ on a

$$\det(uv) = \det u \det v$$

COROLLAIRE 4-4.35 Pour A et $B \in \mathcal{M}_n(\mathbb{K})$

$$\det(AB) = \det A \det B$$

En particulier, les applications "déterminant" sont des morphismes des groupes $(\mathcal{GL}(\mathbb{E}_n), \circ)$ et $(\mathcal{GL}_n(\mathbb{K}), \times)$ dans le groupe multiplicatif $(\mathbb{K}^*, \times)$. Il sont évidemment surjectifs puisqu'il est clair que, pour $\alpha \in \mathbb{K}^*$

$$\det(\text{Diag}(\alpha, 1, \ldots, 1)) = \alpha$$

Si $u \in \mathcal{GL}(\mathbb{E}_n)$ et $M \in \mathcal{GL}_n(\mathbb{K})$ on a

$$\det(u^{-1}) = \frac{1}{\det u} \text{ et } \det(A^{-1}) = \frac{1}{\det A}$$

Chacun des noyaux de ces morphismes est un sous-groupe du groupe de départ, dit "groupe spécial linéaire" (de $\mathbb{E}_n$ ou d'ordre n) :

$$\begin{aligned} \mathcal{SL}(\mathbb{E}_n) &= \{u \in \mathcal{GL}(\mathbb{E}_n) \mid \det u = 1\} \\ \mathcal{SL}_n(\mathbb{K}) &= \{M \in \mathcal{GL}_n(\mathbb{K}) \mid \det M = 1\} \end{aligned}$$

REMARQUE 4-4.36 La multiplication des matrices n'est pas commutative. Cependant, pour calculer le déterminant d'une matrice M qui peut se factoriser sous la forme $M = AB$, il est parfois intéressant de calculer plutôt le déterminant de BA.

4-4.4 Exemples de calculs de déterminants

4-4.4.1 Cas $n = 2$ ou $n = 3$: règle de Sarrus

Pour $n = 2$, il n'y a que deux permutations : l'identité et la transposition $(1, 2)$. On a donc

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Pour $n = 3$, il y a trois permutations impaires : les transpositions $(1,2)$, $(1,3)$ et $(2,3)$. Il y a également 3 permutations paires : l'identité et les deux cycles de longueur 3 notés $(1,2,3)$ et $(1,3,2)$.[2] On a donc

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \delta_1 - \delta_2$$

$$\text{avec } \delta_1 = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$\text{et } \delta_2 = a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32} + a_{13}a_{22}a_{31}$$

On remarque que δ_1 (resp. δ_2) est égal à la somme des produits des trois termes situés sur les trois diagonales "descendantes" (resp. "montantes") du tableau

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

obtenu en écrivant à la droite de la matrice A les deux premières colonnes.
Dès que $n \geqslant 4$, les choses se corsent. Il est d'ailleurs à noter que, dans le cas particulier $n = 3$, il est souvent préférable de d'abord faire apparaître deux zéros sur une ligne ou une colonne, ce qui ramène au calcul d'un déterminant d'ordre 2.

4-4.4.2 Utilisation de la multilinéarité et l'antisymétrie

C'est la technique de base. On la verra à l'oeuvre dans le paragraphe consacré aux manipulations élémentaires. Un exercice un peu plus "théorique" :

EXERCICE 4-4.37 Montrer que, pour tout x réel on a

$$\begin{vmatrix} x^2 & (x+1)^2 & (x+2)^2 & (x+3)^2 \\ (x+1)^2 & (x+2)^2 & (x+3)^2 & (x+4)^2 \\ (x+2)^2 & (x+3)^2 & (x+4)^2 & (x+5)^2 \\ (x+3)^2 & (x+4)^2 & (x+5)^2 & (x+6)^2 \end{vmatrix} = 0$$

(Indication : trouver une base de $\mathbf{R}_2[X]$ adaptée au problème).

[2]Dans S_n, si $\{a_1, a_2, \dots, a_l\}$ est une partie de cardinal l de $\{1, \dots, n\}$, on note $(a_1, a_2, \dots, a_l)$ la permutation circulaire

$$a_1 \to a_2 \to \dots \to a_l \to a_1$$

(permutation qui laisse invariants les éléments de $\{1, \dots, n\} - \{a_1, a_2, \dots, a_l\}$). Une telle permutation a même parité que $l - 1$, car il est facile de la décomposer en produit de transpositions :

$$(a_1, a_2, \dots, a_l) = (a_1, a_l) \cdots (a_1, a_3)(a_1, a_2)$$

(pour vérifier cette égalité, se souvenir que le terme de droite est un composé de transpositions, qui se lit de la droite vers la gauche).

4-4.4.3 Déterminants triangulaires par blocs

Si $T = (t_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ est une matrice triangulaire (supérieure par exemple), le développement de son déterminant par rapport à la première colonne et une récurrence élémentaire donne

$$\det T = \prod_{i=1}^{n} t_{ii}$$

Cette formule se généralise au cas des matrices triangulaires par blocs :

THÉORÈME 4-4.38 Si $A \in \mathcal{M}_p(\mathbb{K})$ et $B \in \mathcal{M}_{n-p}(\mathbb{K})$, quelle que soit la matrice C dans $\mathcal{M}_{p,n-p}(\mathbb{K})$ on a

$$\det \begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = \det A \det B$$

Démonstration : Le calcul est évident si $A = I_p$ ou $B = I_{n-p}$ (développer dans le premier cas par rapport à la première colonne, par rapport à la dernière ligne dans le second). Le cas général découle de la factorisation

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} A & C \\ 0 & I_{n-p} \end{pmatrix}$$

(On pourrait remarquer aussi que, B et C étant fixées, l'application qui aux p vecteurs colonnes de A associe le déterminant à calculer est une forme p-linéaire alternée sur $\mathbb{K}^p$, et est donc proportionnelle au déterminant de A). ∎

On déduit immédiatement, par récurrence sur p :

COROLLAIRE 4-4.39 Si une matrice M est décomposée par blocs (les tailles des différents blocs de colonnes correspondant à celles des blocs de lignes : sur la diagonale apparaissent des blocs carrés)

$$M = \begin{pmatrix} A_{11} & ? & ? & ? & ? \\ 0 & \ddots & ? & & ? \\ \vdots & 0 & A_{ii} & ? & ? \\ \vdots & & 0 & \ddots & ? \\ 0 & \cdots & \cdots & 0 & A_{pp} \end{pmatrix}$$

on a l'égalité

$$\det M = \prod_{i=1}^{p} \det A_{ii}$$

COROLLAIRE 4-4.40 Soit $u \in \mathcal{L}(\mathbb{E})$, avec $\mathbb{E}$ de dimension finie décomposé en somme directe de sous-espaces stables par u

$$\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{F}_i \text{ avec } u(\mathbb{F}_i) \subset \mathbb{F}_i$$

Si on note u_i l'endomorphisme de $\mathbb{F}_i$ induit par u, on a

$$\det u = \prod_{i=1}^{p} \det u_i$$

Démonstration : Il suffit d'écrire la matrice de u dans une base adaptée à la décomposition $\mathbb{E} = \mathbb{F}_1 \oplus \cdots \oplus \mathbb{F}_p$. ■

4-4.4.4 Utilisation de polynômes

De nombreux exercices font intervenir des raisonnements sur les polynômes. L'exemple le plus classique est celui du déterminant de Vandermonde :

PROPOSITION 4-4.41 Si $(x_i)_{1\leqslant i\leqslant n}$ sont n scalaires, on appelle déterminant de Vandermonde associé à ces scalaires le déterminant :

$$V(x_1,\ldots,x_n) = \begin{vmatrix} 1 & \cdots & 1 & \cdots & 1 \\ x_1 & & x_j & & x_n \\ x_1^2 & & x_j^2 & & x_n^2 \\ \vdots & & \vdots & & \vdots \\ x_1^{n-1} & \cdots & x_j^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$

On a alors

$$V(x_1,\ldots,x_n) = \prod_{1\leqslant i<j\leqslant n} (x_j - x_i)$$

Démonstration : On peut d'abord remarquer que la formule est évidente lorsqu'au moins deux des x_i sont égaux. Dans le cas général, on peut remarquer qu'ajouter à la dernière ligne une combinaison linéaire des autres lignes fait apparaître en dernière ligne les valeurs d'un polynôme P de degré $n-1$ et de coefficient dominant égal à 1 sur les différents x_i. Il est facile de trouver un tel polynôme pour que la dernière ligne commence alors par $n-1$ zéros. On prend

$$P = \prod_{i=1}^{n-1} (X - x_i) = X^{n-1} - \sigma_1 X^{n-2} + \cdots + (-1)^{n-1} \sigma_{n-1}$$

où les σ_i sont les fonctions symétriques élémentaires[3] des $(x_i)_{1\leqslant i\leqslant n-1}$. En effectuant la manipulation

$$L_n \longleftarrow L_n - \sigma_1 L_{n-1} + \cdots + (-1)^{n-1} \sigma_{n-1} L_1$$

[3]Rappelons que l'on a notamment

$$\sigma_1 = x_1 \cdots + x_{n-1},\ \sigma_2 = \sum_{1\leqslant i<j\leqslant n-1} x_i x_j \text{ et enfin } \sigma_{n-1} = x_1 \cdots x_{n-1}$$

et en développant par rapport à la dernière ligne, on obtient alors

$$V(x_1, \ldots, x_n) = P(x_n) V(x_1, \ldots, x_{n-1})$$

ce qui permet de conclure aisément par récurrence. ■

EXERCICE 4-4.42 Par un raisonnement analogue, calculer le déterminant de Cauchy

$$\Delta(a_1, \ldots, a_n, b_1, \ldots, b_n) = \det\left(\frac{1}{a_i + b_j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

lorsque les $2n$ scalaires a_i et b_j vérifient

$$\forall i, j \quad a_i + b_j \neq 0$$

(on utilisera ici la notion de décomposition d'une fraction rationnelle en éléments simples).

EXERCICE 4-4.43 Calculer

$$\begin{vmatrix} 1 & \cos a_1 & \cos 2a_1 & \cdots & \cos(n-1)a_1 \\ 1 & \cos a_2 & \cos 2a_2 & \cdots & \cos(n-1)a_2 \\ \vdots & & & & \\ 1 & \cos a_n & \cos 2a_n & \cdots & \cos(n-1)a_n \end{vmatrix}$$

Indication : $\cos kx$ est un polynôme de degré k en $\cos x$. Quel est son coefficient dominant ?

Terminons par un exercice très classique :

EXERCICE 4-4.44 Calculer $\det A$ avec

$$A = \begin{pmatrix} x_1 & b & b & b & b \\ a & \ddots & b & & b \\ \vdots & a & x_i & b & b \\ \vdots & & a & \ddots & b \\ a & \cdots & \cdots & a & x_n \end{pmatrix}$$

Lorsque $a = b$, ce n'est pas très difficile. Si a ou b est nul, encore moins. Si J est la matrice carrée d'ordre n dont tous les coefficients sont égaux à 1, l'application

$$x \mapsto \det(A + xJ)$$

est polynomiale. Quel est le degré du polynôme qu'on peut lui associer ?

4-4.4.5 Dérivée d'un déterminant

Nous reviendrons dans le chapitre sur la dérivation sur le résultat suivant :

EXERCICE 4-4.45 Si $A(t) = (a_{ij}(t))_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ est une matrice à coefficients complexes dont les coefficients sont des fonctions d'une variable réelle t dérivables sur un intervalle I, l'application φ

$$\varphi : t \mapsto \det A(t)$$

est dérivable sur I et $\varphi'(t)$ est somme de n déterminants,

$$\varphi'(t) = \sum_{i=1}^{n} \det A_i(t)$$

où $A_i(t)$ est la matrice déduite de $A(t)$ en remplaçant la $i^{\text{ème}}$ ligne par sa dérivée au point t. (Indication : quelle est la dérivée de $t \mapsto a_{1\sigma(1)}(t)\, a_{2\sigma(2)}(t) \cdots a_{n\sigma(n)}(t)$?). On a un résultat analogue en travaillant sur les colonnes.

EXEMPLE 4-4.46 Calculer

$$D_n(x) = \begin{vmatrix} x & 1 & 0 & \cdots & 0 \\ \frac{x^2}{2!} & x & 1 & \ddots & \vdots \\ \vdots & \frac{x^2}{2!} & x & 1 & 0 \\ \frac{x^{n-1}}{(n-1)!} & \vdots & \frac{x^2}{2!} & x & 1 \\ \frac{x^n}{n!} & \frac{x^{n-1}}{(n-1)!} & \cdots & \frac{x^2}{2!} & x \end{vmatrix}$$

4-4.4.6 Utilisation de produits matriciels ou de valeurs propres

EXEMPLE 4-4.47 Déterminant circulant. Soit M la matrice de $\mathcal{M}_n(\mathbb{C})$

$$M = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_n & a_1 & a_2 & \cdots & \vdots \\ a_{n-1} & \ddots & \ddots & \ddots & a_3 \\ \vdots & \ddots & \ddots & \ddots & a_2 \\ a_2 & a_3 & \cdots & a_n & a_1 \end{pmatrix}$$

Pour calculer le déterminant de M, on calcule le produit $M\Omega$, où Ω est la transposée de la matrice de Vandermonde associée aux scalaires

$$(1, \omega, \omega^2, \ldots, \omega^{n-1})$$

avec $\omega = e^{\frac{2i\pi}{n}}$ (c'est-à-dire aux racines complexes $n^{\text{ièmes}}$ de l'unité). Qu'observe-t-on et pourquoi a-t-on multiplié par Ω ? Parce que ses vecteurs colonnes sont *vecteurs propres de* M. On s'aperçoit sur cet exemple que le déterminant de M est égal au produit des *valeurs propres de* M. Nous y reviendrons dans le chapitre sur la réduction des endomorphismes.

4-4.5 Applications des déterminants

Répétons qu'il s'agit essentiellement d'applications à usage théorique. Il est par exemple très rare qu'on utilise les déterminants pour calculer effectivement le rang d'une matrice.

4-4.5.1 Rang d'un système de vecteurs, d'une matrice

Le théorème **4-1.**32 donne immédiatement :

THÉORÈME 4-4.48 Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ **est non nulle, on a** $\operatorname{rang} A = r$ **si et seulement si il existe un déterminant d'ordre** r **extrait de** A **non nul, tout déterminant d'ordre** $r+1$ **étant nul (s'il en existe).**

Pour étudier le rang d'un système de vecteurs d'un espace de dimension finie, on pourra travailler avec la matrice représentative de ces vecteurs dans une base arbitraire de l'espace.

EXERCICE 4-4.49 On dit qu'une suite de matrices $A_k = (a_{ij}(k))_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ de $\mathcal{M}_n(\mathbb{C})$ est convergente si les n^2 suites complexes $(a_{ij}(k))_{k \in \mathbb{N}}$ sont convergentes. On pose alors

$$\lim_{k \to +\infty} A_k = \left(\lim_{k \to +\infty} a_{ij}(k) \right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

Montrer que la limite d'une suite de matrices de rang r est de rang inférieur ou égal à r. Montrer qu'on peut avoir inégalité stricte.

4-4.5.2 Formules de Cramer

Il s'agit d'une autre façon de présenter le théorème **4-4.**22. On considère un système linéaire de n équations à n inconnues (scalaires) :

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \qquad \vdots & \quad \vdots \\ a_{i1}x_1 + \cdots + a_{in}x_n &= b_i \\ \qquad \vdots & \quad \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n \end{cases}$$

qu'on peut écrire sous forme matricielle

$$AX = B \quad \text{avec} \quad A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(\mathbb{K}) \ \text{ et } B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$$

Ce système est dit de Cramer si et seulement si la matrice A est inversible. Quel que soit B, le système possède une solution unique

$$X = A^{-1}B$$

Les déterminants permettent d'expliciter cette solution. On note $\Delta = \det A$ (le "déterminant du système") et on note Δ_j le déterminant de la matrice obtenue en remplaçant la $j^{\text{ème}}$ colonne de A (celle correspondant aux coefficients de l'inconnue x_j dans le système) par la colonne B (les termes "du second membre"). Si $(x_1, \ldots, x_n)$ est la solution du système, en notant C_k la $k^{\text{ème}}$ colonne de A

$$\Delta_j = \det\left(C_1, \ldots, \underset{j^{\text{ème}}\text{colonne}}{\underset{\downarrow}{\sum_{i=1}^{n} x_i C_i}}, \ldots, C_n\right) = x_j \det(C_1, \ldots, C_j, \ldots, C_n) = x_j \Delta$$

et on obtient les "formules de Cramer" :

$$\boxed{x_j = \frac{\Delta_j}{\Delta}}$$

En développant le déterminant Δ_j selon sa $j^{\text{ème}}$ colonne et en écrivant la solution du système sous forme d'une égalité matricielle

$$X = CB$$

on retrouve l'expression de $A^{-1} = C$ à l'aide de la comatrice de A.

4-4.5.3 Equations d'un sous-espace vectoriel

On travaille sur un espace vectoriel $\mathbb{E}_n$ rapporté à une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$. Pour caractériser un espace vectoriel $\mathbb{F}_p$ de dimension p, on peut s'en donner une base $\mathcal{B}' = (u_j)_{1 \leqslant j \leqslant p}$, représentée par sa matrice (de rang p) dans la base $\mathcal{B}$

$$A = M_{\mathcal{B}}(\mathcal{B}') = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$$

On sait aussi (théorème **3-3**.5) que $\mathbb{F}_p$ peut s'exprimer comme intersection de noyaux de $n - p$ formes linéaires indépendantes (correspondant à un "système d'équations" de $\mathbb{F}_p$). De telles formes linéaires peuvent se calculer en utilisant les déterminants : comme A est de rang p, on peut en extraire p lignes indépendantes. En renumérotant éventuellement les vecteurs de $\mathcal{B}$, on peut supposer que la matrice

$$A' = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant p}} \in \mathcal{GL}_p(\mathbb{K})$$

Si $x \in \mathbb{E}_n$ est représenté dans la base $\mathcal{B}$ par la colonne X de ses coordonnées, dire que x est dans $\mathbb{F}_p$ revient à dire que X est combinaison linéaire des vecteurs colonnes de A c'est-à-dire que la matrice de $\mathcal{M}_{n,p+1}(\mathbb{K})$ décomposée en blocs $B = (A \mid X)$ est encore de rang p. En particulier, les $n - p$ déterminants d'ordre $p + 1$ extraits de B

$$\Delta_i(X) = \begin{vmatrix} & & & x_1 \\ & \boxed{A'} & & \vdots \\ & & & x_p \\ a_{i1} & \cdots & a_{ip} & x_i \end{vmatrix} \quad \text{pour } i = p+1, \ldots, n \quad \text{où } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(obtenus en "bordant" A' par les coordonnées d'indices correspondant de X à droite, et par une ligne composée d'une ligne supplémentaire de A et de la coordonnée de X correspondante) sont tous nuls.

Il est facile de voir que les $n-p$ applications

$$x \mapsto \Delta_i(X)$$

sont des formes linéaires indépendantes sur $\mathbb{E}_n$: en développant $\Delta_i(X)$ par rapport à sa dernière colonne, on s'aperçoit que Δ_i est combinaison linéaire des formes coordonnées $(\varphi_k)_{1\leqslant k\leqslant n}$ dans la base $\mathcal{B}$. Comme de plus la coordonnée de Δ_i selon φ_i est le cofacteur de x_i soit $\det A' \neq 0$ et que, pour $j \neq i$ choisis dans $\{p+1,\ldots,n\}$, Δ_j n'a pas de composante selon φ_i, l'indépendance de ces formes n'est pas difficile à prouver.

On obtient donc $n-p$ formes linéaires indépendantes identiquement nulles sur $\mathbb{F}_p$. Elles donnent un système d'équations de $\mathbb{F}_p$:

$$x \in \mathbb{F}_p \Leftrightarrow \forall i \in \{p+1,\ldots,n\} \quad \Delta_i(X) = 0$$

EXEMPLE 4-4.50 Equations dans $\mathbb{R}^4$ de $\mathbb{V} = \text{vect}\{(1,2,3,4),(2,4,0,1)\}$. On a ici, en travaillant dans la base canonique de $\mathbb{R}^4$,

$$A = \begin{pmatrix} \boxed{1} & \boxed{2} \\ 2 & 4 \\ \boxed{3} & \boxed{0} \\ 4 & 1 \end{pmatrix}$$

On a souligné une matrice d'ordre 2 inversible extraite de A. Un vecteur de $\mathbb{R}^4$ $X = (x,y,z,t)$ est dans $\mathbb{V}$ si et seulement si

$$\begin{vmatrix} 1 & 2 & x \\ 3 & 0 & z \\ 2 & 4 & y \end{vmatrix} = \begin{vmatrix} 1 & 2 & x \\ 3 & 0 & z \\ 4 & 1 & t \end{vmatrix} = 0$$

soit $12x - 6y = 7z - 6t + 3x = 0$. On pourra donc prendre par exemple

$$X \in \mathbb{V} \Leftrightarrow \begin{cases} 2x - y = 0 \\ 7z - 6t + 3x = 0 \end{cases}$$

Nous reviendrons dans la section suivante sur une utilisation possible de ces déterminants pour trouver des conditions de compatibilité pour un système linéaire.

4-5 Application aux systèmes d'équations linéaires

Dans toute cette section, on considère un système de p équations linéaires scalaires à n inconnues $(x_1,\ldots,x_n)$:

$$(S) : \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1j}x_j + \cdots + \alpha_{1n}x_n &= \beta_1 \\ \vdots & \vdots \\ \alpha_{i1}x_1 + \cdots + \alpha_{ij}x_j + \cdots + \alpha_{in}x_n &= \beta_i \\ \vdots & \vdots \\ \alpha_{p1}x_1 + \cdots + \alpha_{pj}x_j + \cdots + \alpha_{pn}x_n &= \beta_p \end{cases}$$

4-5.1 Trois interprétations possibles d'un système

La matrice des coefficients

$$A = (\alpha_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$$

est dite matrice du système. Son rang est, par définition, le rang du système. Nous donnons ici diverses interprétations de ce système. Avec un point de vue différent, on arrive aux mêmes résultats :

- **Une équation linéaire (vectorielle)** : Si on note Φ_A le morphisme de $\mathbb{K}^n$ vers $\mathbb{K}^p$ canoniquement associé à A, et X la colonne des inconnues , on résout

$$\Phi_A(X) = B, \text{ avec } B = \begin{pmatrix} \beta_1 \\ \cdots \\ \beta_i \\ \cdots \\ \beta_p \end{pmatrix}$$

  Le système possède des solutions si $B \in \operatorname{Im} \Phi_A$, qui est un sous-espace de dimension r de $\mathbb{K}^p$, et qui est donc caractérisé par un système de $p - r$ équations. Si B vérifie ces $p - r$ équations (on parle de **$p - r$ conditions de compatibilité**) le théorème du rang et la section **2-1.7** montrent que **l'ensemble des solutions du système est un sous-espace affine de $\mathbb{K}^n$, de dimension $n - r$.**

- **Ecriture d'un vecteur comme combinaison linéaire d'une famille** : Si on note $U_1, \ldots, U_n$ les vecteurs colonnes de la matrice A (formant donc un système de rang r dans $\mathbb{K}^p$), le système peut s'écrire

$$x_1 U_1 + \cdots + x_n U_n = B$$

  (il n'y a, à vrai dire, pas grande différence avec l'interprétation précédente : $U_1, \ldots, U_n$ forment une base de $\operatorname{Im} \Phi_A$). On cherche donc à exprimer B comme combinaison linéaire de $U_1, \ldots, U_n$. Les conditions de compatibilité sont encore ici

$$B \in \operatorname{vect}(U_1, \ldots, U_n)$$

  soit $p - r$ conditions, puisque cet espace est de dimension r. Si elles sont vérifiées, (en supposant, pour simplifier l'écriture, que $U_1, \ldots, U_r$ sont indépendants), le système s'écrit

$$x_1 U_1 + \cdots + x_r U_r = B - \sum_{j=r+1}^{n} x_j U_j$$

  Quelles que soient les valeurs de $(x_j)_{r+1 \leqslant j \leqslant n}$ (dites "inconnues secondaires"), les "inconnues principales" $(x_i)_{1 \leqslant i \leqslant r}$ s'expriment en fonction de ces inconnues secondaires : x_i est la coordonnée selon U_i du vecteur $B - \sum_{j=r+1}^{n} x_j U_j$

(qui appartient à vect $(U_1, \ldots, U_r)$). Un instant de réflexion montre alors que les $(x_i)_{1 \leqslant i \leqslant r}$ sont des fonctions affines des $(x_j)_{r+1 \leqslant j \leqslant n}$, ce qui doit permettre de se convaincre que l'ensemble des solutions est un sous-espace affine de $\mathbb{K}^n$ de dimension $n - r$.

- **Interprétation par les formes linéaires** : C'est ici la théorie de la dualité qui intervient. Chaque ligne de la matrice A peut être interprétée comme définissant une forme linéaire sur $\mathbb{K}^n$ (on identifie en fait la forme et sa matrice dans la base canonique). Le système s'écrit alors

$$\begin{cases} L_1(X) = \beta_1 \\ \vdots \\ L_p(X) = \beta_p \end{cases}$$

Le rang de la famille $(L_i)_{1 \leqslant i \leqslant p}$ est r. On peut alors extraire une sous-famille libre de cardinal r. Quitte à renuméroter les équations, on peut supposer que c'est la famille $(L_i)_{1 \leqslant i \leqslant r}$. Les r équations correspondantes sont alors considérées comme "principales". Les conditions de compatibilité sautent alors aux yeux :

Si $k \in \{r+1, \ldots, p\}$ (soit $p - r$ valeurs), la forme L_k est combinaison linéaire des $(L_i)_{1 \leqslant i \leqslant r}$, soit

$$L_k = \sum_{i=1}^{r} \mu_{ik} L_i$$

Pour que le système possède des solutions, il est évidemment nécessaire que les "seconds membres" vérifient les mêmes relations :

$$\beta_k = \sum_{i=1}^{r} \mu_{ik} \beta_i$$

Si ces $p - r$ égalités sont vérifiées, les solutions du système sont alors clairement celles du système des équations principales (on dit que le système d'équations principales est équivalent au système de départ)

$$\begin{cases} L_1(X) = \beta_1 \\ \vdots \\ L_r(X) = \beta_r \end{cases}$$

Le théorème **3-2**.1 montre alors que ce système possède des solutions, qui forment un sous-espace affine de $\mathbb{K}^n$ de dimension $n - r$.

Résumons les discussions précédentes :

THÉORÈME 4-5.1 Pour qu' un système linéaire de p équations à n inconnues et de rang r soit compatible (c'est-à-dire possède au moins une solution), il y a $p - r$ conditions à vérifier. Pour ce faire, on peut extraire un système de r équations indépendantes et vérifier la compatibilité des $p - r$ équations restantes avec ce système. Lorsqu'il y a compatibilité, le système est alors équivalent au système d'équations principales et l'ensemble des solutions est un espace affine de dimension $n - r$.

L'algorithme du pivot de Gauss permettra de redémontrer ce résultat (c'est en fait un quatrième point de vue, celui du calcul "numérique" des solutions). Contentons nous pour le moment de montrer comment les déterminants permettraient d'écrire (théoriquement) les solutions.

4-5.2 Utilisation de déterminants

Dire que la matrice du système est de rang r permet de considérer un déterminant extrait d'ordre r non nul. Quitte à renuméroter les équations et les inconnues, on peut supposer

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} \end{vmatrix} \neq 0$$

Les r premières équations sont alors indépendantes. On les décrète "principales" tout comme les r premières inconnues. Les colonnes correspondantes $U_1, \ldots, U_r$ sont indépendantes, et les conditions de compatibilité (qui se résument comme on l'a vu à $B \in \text{vect}(U_1, \ldots, U_r)$) peuvent s'écrire (*cf.* section **4-4.5.3**)

$$\forall i \in \{r+1, \ldots, p\} \quad \Delta_i = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1r} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \beta_r \\ \alpha_{i1} & \cdots & \alpha_{ir} & \beta_i \end{vmatrix} = 0$$

Si ces conditions sont vérifiées, on peut résoudre le système des r premières équations comme un système de Cramer, en faisant passer les inconnues "secondaires" au second membre :

$$(S) \Leftrightarrow \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1r}x_r &= \beta_1 - \sum_{k=r+1}^{n} \alpha_{1k}x_k \\ \vdots &= \vdots \\ \alpha_{i1}x_1 + \cdots + \alpha_{ir}x_r &= \beta_i - \sum_{k=r+1}^{n} \alpha_{ik}x_k \\ \vdots &= \vdots \\ \alpha_{r1}x_1 + \cdots + \alpha_{rr}x_r &= \beta_r - \sum_{k=r+1}^{n} \alpha_{rk}x_k \end{cases}$$

ce qui donne, pour $j \in \{1, \ldots, r\}$ l'expression (peu utilisable en pratique)

$$x_j = \frac{D_j}{D} - \sum_{k=r+1}^{n} x_k \frac{D_j(k)}{D}$$

où D est le déterminant $\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} \end{vmatrix}$, D_j (resp. $D_j(k)$) le déterminant qui s'en déduit en remplaçant la $j^{\text{ème}}$ colonne de D par la colonne $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}$ (respectivement par la colonne $\begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{rk} \end{pmatrix}$).

EXEMPLE 4-5.2 Si a, b et c sont trois constantes réelles, discuter et résoudre le système

$$\begin{cases} (b+c)\,x + (bc-1)\,y + (1+b^2)\,(1+c^2)\,z = a \\ (c+a)\,x + (ca-1)\,y + (1+c^2)\,(1+a^2)\,z = b \\ (a+b)\,x + (ab-1)\,y + (1+a^2)\,(1+b^2)\,z = c \end{cases}$$

Il est peut-être un cas particulier important à envisager : celui de la résolution d'un système **homogène** de $n-1$ équations indépendantes à n inconnues. On sait alors que l'ensemble des solutions du système est une droite vectorielle de $\mathbb{K}^n$. Il suffit donc de trouver une solution non nulle. Pour ce faire, on résout le système de Cramer obtenu en rajoutant une $n^{\text{ième}}$ équation fictive supposée indépendante des précédentes :

$$\begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1j}x_j + \cdots + \alpha_{1n}x_n & = & 0 \\ & \vdots & \\ \alpha_{n-1,1}x_1 + \cdots + \alpha_{n-1,j}x_j + \cdots + \alpha_{n-1,n}x_n & = & 0 \\ \alpha_{n1}x_1 + \cdots + \alpha_{nj}x_j + \cdots + \alpha_{nn}x_n & = & \lambda \end{cases}$$

Si Δ est le déterminant de ce système, les formules de Cramer donnent

$$x_j = \lambda \frac{C_{nj}}{\Delta}$$

où C_{nj} est le cofacteur de α_{nj} dans la matrice du système.

Les solutions du système de départ sont donc proportionnelles aux cofacteurs des coefficients d'une ligne fictive ajoutée à la matrice du système. (On peut faire le rapprochement avec le raisonnement qu'on fait, en utilisant la notion de produit vectoriel, pour déterminer un vecteur directeur de la droite d'équations

$$\begin{cases} x + y + z = 3 \\ 2x - 3y + 4z = 1 \end{cases}$$

dans un repère **orthonormé** d'un espace affine euclidien de dimension 3).

4-6 Opérations élémentaires

Nous avons jusqu'à présent privilégié l'aspect vectoriel, géométrique, pour présenter le calcul matriciel. Un autre point de vue, plus "numérique", est tout aussi important, et permet de retrouver certains des résultats étudiés précédemment.

4-6.1 Matrices élémentaires

Il s'agit de matrices carrées inversibles, de trois types différents. On travaille dans $\mathcal{M}_n(\mathbb{K})$, dont la base canonique est notée $(E_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$

- <u>**Matrice de type $\mathcal{T}_1$**</u> : Si i et j sont **distincts** dans $\{1,\dots,n\}$, et $\lambda\in\mathbb{K}$, on note

$$\boxed{M_{ij}(\lambda)=I_n+\lambda E_{ij}}$$

C'est une matrice avec une diagonale de 1, et λ à l'intersection de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne, tous les autres coefficients étant nuls. Ces matrices sont aussi dites matrices élémentaires de transvection, pour une raison qui apparaîtra plus loin.

Pour i et j fixés, l'application

$$(\mathbb{K},+)\to(\mathcal{GL}_n(\mathbb{K}),\times)\quad\lambda\mapsto M_{ij}(\lambda)$$

est un morphisme de groupes, c'est-à-dire $M_{ij}(\lambda)\,M_{ij}(\mu)=M_{ij}(\lambda+\mu)$.

- <u>**Matrice de type $\mathcal{T}_2$**</u> : Si $i\in\{1,\dots,n\}$ et $\alpha\in\mathbb{K}^*$, on note

$$D_i(\alpha)=I_n+(\alpha-1)\,E_{ii}=\text{Diag}\Big(1,\dots,1,\underset{i^{\text{ème}}\text{ place}}{\alpha},1,\dots,1\Big)$$

Cette fois, c'est l'application

$$(\mathbb{K}^*,\times)\to(\mathcal{GL}_n(\mathbb{K}),\times)\quad\alpha\mapsto D_i(\alpha)$$

qui est un morphisme de groupes : $D_i(\alpha)\,D_i(\beta)=D_i(\alpha\beta)$. Ces matrices sont aussi appelées matrices élémentaires de dilatation.

- <u>**Matrice de type $\mathcal{T}_3$**</u> : On les appelle aussi matrices de permutation. Si $\sigma\in(S_n,\circ)$, on note P_σ la matrice de passage de la base canonique $(e_i)_{1\leqslant i\leqslant n}$ de $\mathbb{K}^n$ à la base $(e_{\sigma(i)})_{1\leqslant i\leqslant n}$. Le morphisme de $\mathbb{K}^n$ canoniquement associé à P_σ est l'unique automorphisme f_σ de $\mathbb{K}^n$ vérifiant

$$e_i\mapsto f_\sigma(e_i)=e_{\sigma(i)}.$$

L'application

$$(S_n,\circ)\to(\mathcal{GL}(\mathbb{K}^n),\circ)\ \text{ définie par }\sigma\mapsto f_\sigma$$

est clairement un morphisme de groupes. Il en est donc de même de l'application

$$(S_n, \circ) \to (\mathcal{GL}_n(\mathbb{K}), \times) \qquad \sigma \mapsto P_\sigma.$$

En particulier on a $(P_\sigma)^{-1} = P_{\sigma^{-1}}$.

Dans chaque colonne de P_σ, il y a $n-1$ zéros et un 1. Plus précisément, à la $j^{\text{ème}}$ colonne, le coefficient a_{ij} est nul sauf si $i = \sigma(j)$, auquel cas il est égal à 1 :

$$a_{ij} = 1 \Leftrightarrow i = \sigma(j) \Leftrightarrow j = \sigma^{-1}(j)$$

Il en résulte clairement que l'on a aussi

$$(P_\sigma)^{-1} = P_{\sigma^{-1}} = {}^tP_\sigma$$

Dans chaque ligne de P_σ il y a également $n-1$ zéros et un 1. Lorsque le corps de base est $\mathbb{R}$, la matrice P_σ peut être considérée comme matrice orthogonale (voir le chapitre sur les espaces euclidiens). Comme le groupe S_n est engendré par les transpositions, nous serons dans la suite amenés à considérer le cas particulier des matrices de transposition : si $\tau = (i,j)$, la matrice P_τ est la matrice qui se déduit de I_n en permutant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ ligne. C'est une matrice symétrique, égale à son inverse (c'est la matrice, dans une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ d'un ev, de la symétrie par rapport à l'hyperplan

$$\text{vect}(e_k)_{k \notin \{i,j\}} \oplus \text{vect}(e_i + e_j)$$

parallèlement à $\text{vect}(e_i - e_j)$).

Evidemment, si une permutation σ se décompose en produit de k transpositions

$$\sigma = \tau_1 \cdots \tau_k$$

on aura également

$$P_\sigma = P_{\tau_1} \cdots P_{\tau_k}$$

4-6.2 Opérations élémentaires sur les lignes ou les colonnes d'une matrice

Si $A \in \mathcal{M}_{p,n}(\mathbb{K})$, on définit trois types de manipulations élémentaires sur les lignes de A, qui correspondent à une prémultiplication par une matrice élémentaire de $\mathcal{GL}_p(\mathbb{K})$:

- **Opérations de type $\mathcal{L}_1$** : si i et j sont deux indices différents et $\lambda \in \mathbb{K}^*$, on ajoute à la $i^{\text{ème}}$ ligne de A la $j^{\text{ème}}$ ligne multipliée par λ. Si A est considérée comme matrice de vecteurs lignes $(L_i)_{1 \leqslant i \leqslant p}$, on symbolise cette opération par

$$L_i \leftarrow L_i + \lambda L_j$$

L'interprétation de la prémultiplication de A par une matrice en termes de manipulations sur les lignes de A montre que cette opération revient à remplacer A par la matrice $M_{ij}(\lambda)\,A$, ce que l'on écrira en abrégé

$$A \leftarrow M_{ij}(\lambda)\,A$$

Il est clair que l'opération "réciproque" est simplement $A \leftarrow M_{ij}(-\lambda)\,A$.

- **Opérations de type $\mathcal{L}_2$** : On multiplie une ligne de A par un scalaire non nul. Si $\alpha \in \mathbb{K}^*$, on symbolise par

$$L_i \leftarrow \alpha L_i$$

ce qui correspond évidemment à

$$A \leftarrow D_i(\alpha)\,A$$

- **Opérations de type $\mathcal{L}_3$** : On permute deux lignes de A : si $i \neq j$, on écrira

$$L_i \leftrightarrow L_j$$

cette opération qui correspond en fait, si $\tau = (i, j)$, à

$$A \leftarrow P_\tau A$$

On peut être parfois amené à considérer plusieurs opérations successives de ce type. Si $\sigma = \tau_1 \cdots \tau_k$, la prémultiplication de A par P_σ correspond en fait à effectuer d'abord la transposition τ_k sur les lignes de A, puis τ_{k-1} etc... et enfin τ_1.

$$A \leftarrow P_\sigma A \quad \text{transforme } A = \begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix} \text{ en } \begin{pmatrix} L_{\sigma^{-1}(1)} \\ \vdots \\ L_{\sigma^{-1}(p)} \end{pmatrix}$$

Comme les matrices élémentaires prémultiplicatrices sont dans $\mathcal{GL}_p(\mathbb{K})$, ces manipulations transforment A en une matrice équivalente.

On définit de la même manière les opérations élémentaires sur les colonnes de la matrice A. Elles correspondent cette fois à une post-multiplication par une matrice élémentaire de $\mathcal{GL}_n(\mathbb{K})$:

$\mathcal{C}_1$: $C_i \leftarrow C_i + \lambda C_j$ correspond à $A \leftarrow AM_{ji}(\lambda)$: attention aux indices !

$\mathcal{C}_2$: $C_i \leftarrow \alpha C_i$ correspond à $A \leftarrow AD_i(\alpha)$ avec bien sûr $\alpha \in \mathbb{K}^*$

$\mathcal{C}_3$: $C_i \leftrightarrow C_j$ correspond à $A \leftarrow AP_\tau$ où τ est la transposition (i, j)

Il est à noter que, cette fois, la post-multiplication par une matrice P_σ avec $\sigma \in S_n$ remplace la matrice $A = (C_1, \ldots, C_n)$ par la matrice $\left(C_{\sigma(1)}, \cdots, C_{\sigma(n)}\right)$.

4-6.3 Lemme du pivot de Gauss

THÉORÈME 4-6.1 Soit $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$. **Si** $a_{1,1} \neq 0$ **il existe une suite de manipulations élémentaires de type** $\mathcal{L}_1$ **sur les lignes de** A **qui transforme** A **en une matrice de la forme**

$$\begin{pmatrix} \alpha_{1,1} & \\ 0 & ? \\ \vdots & \\ 0 & \end{pmatrix}$$

Si $a_{1,1} = 0$ **mais si la première colonne de** A **n'est pas nulle, on arrive au même résultat après permutation de deux lignes de** A.

Démonstration : Si $a_{1,1} \neq 0$, il suffit d'effectuer les manipulations

$$L_i \longleftarrow L_i - \frac{a_{i,1}}{a_{1,1}} L_1 \quad \text{pour } 2 \leqslant i \leqslant p$$

Si $a_{1,1} = 0$ avec $a_{k,1} \neq 0$, il suffit d'opérer d'abord la manipulation $L_1 \leftrightarrow L_k$ pour être ramené à la situation précédente. ■

COROLLAIRE 4-6.2 Si A **est de rang** $r > 0$, **il existe une suite de manipulations** $\mathcal{L}_1$ **et** $\mathcal{L}_3$ **et de type** $\mathcal{C}_3$ **transformant** A **en une matrice de la forme**

$$\left(\begin{array}{|ccc|c} \hline \alpha_{1,1} & ? & ? & \\ 0 & \ddots & ? & ? \\ 0 & 0 & \alpha_{r,r} & \\ \hline 0 & \cdots & 0 & 0_{p-r,n-r} \end{array}\right)$$

où $\alpha_{i,i} \neq 0$ **pour** $1 \leqslant i \leqslant r$. **Cette propriété caractérise les matrices de rang** r. **En particulier,** $A \in \mathcal{M}_n(\mathbb{K})$ **est inversible si et seulement s'il existe une suite de manipulations** $\mathcal{L}_1$ **et** $\mathcal{L}_3$ **transformant** A **en une matrice triangulaire supérieure à coefficients diagonaux tous non nuls.**

Démonstration : Comme $A \neq 0$, il existe une colonne de A non nulle. Après permutation éventuelle avec la première colonne , on se ramène aux hypothèses du théorème précédent, et des manipulations $\mathcal{L}_1$ (et éventuellement une permutation $\mathcal{L}_3$) transforment A en une matrice

$$A' = \begin{pmatrix} \alpha_{1,1} & ? & \cdots & ? \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

avec $\alpha_{1,1} \neq 0$ et $A_1 \in \mathcal{M}_{p-1,n-1}(\mathbb{K})$. Comme A' est de rang r, il est clair que A_1 est de rang $r-1$ (compter le nombre de lignes indépendantes de A'). Il suffit ensuite d'opérer sur les lignes et colonnes

de A_1 et de réitérer le raisonnement (les seules opérations sur les colonnes étant des permutations, on opérera évidemment les mêmes permutations sur les éléments de la première ligne de A' : on pourrait formaliser le raisonnement en procédant par récurrence sur r) pour transformer finalement A en une matrice

$$A'' = \left(\begin{array}{ccc|c} \alpha_{1,1} & ? & ? & \\ 0 & \ddots & ? & ? \\ 0 & 0 & \alpha_{r,r} & \\ \hline 0 & \cdots & 0 & A_r \end{array} \right)$$

avec $\alpha_{i,i} \neq 0$ pour $1 \leqslant i \leqslant r$. Comme il est clair que

$$\operatorname{rang} A'' = r = r + \operatorname{rang} A_r$$

on a forcément $A_r = 0$. Comme les manipulations élémentaires conservent le rang, on a ainsi caractérisé les matrices de rang r. Les matrices carrées inversibles ne sont qu'un cas particulier de ce qui précède. ■

4-6.4 Applications

Outre la détermination pratique du rang d'une matrice obtenue au corollaire **4-6**.2, une démarche basée sur les manipulations élémentaires permet de retrouver des résultats obtenus précédemment par des considérations plus géométriques :

4-6.4.1 Résolution d'un système linéaire d'équations

Les notations sont celles de la section **4-5**. Commençons par l'étude théorique du système

$$AX = B$$

où $A \in \mathcal{M}_{p,n}(\mathbb{K})$ est supposée de rang r. On forme alors **la matrice complète du système**

$$A' = \begin{pmatrix} A & B \end{pmatrix} \in \mathcal{M}_{p,n+1}(\mathbb{K})$$

Chaque ligne de cette matrice correspond à une équation du système de départ, et les manipulations élémentaires sur les lignes de A' transforment cette matrice en matrice complète d'un système d'équations équivalent au système de départ. On se permettra aussi des permutations sur les n premières colonnes de A', ce qui correspond clairement à une renumérotation des inconnues du système. Si on met en oeuvre sur A' les manipulations transformant A comme en **4-6**.2, on obtient un système équivalent dont la matrice complète est de la forme (on suppose ici $r < \min(p,n)$, les cas $r = p$ ou $r = n$ s'en déduisant avec une écriture légèrement

différente)

$$A'' = \begin{pmatrix} \gamma_{1,1} & \cdots & \gamma_{1,r} & \gamma_{1,r+1} & \cdots & \gamma_{1,n} & b'_1 \\ 0 & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \gamma_{r,r} & \gamma_{r,r+1} & \cdots & \gamma_{r,n} & b'_r \\ & & & & & & b'_{r+1} \\ & 0 & & & 0 & & \vdots \\ & & & & & & b'_p \end{pmatrix}$$

avec $\gamma_{i,i} \neq 0$ pour $i = 1, \ldots, r$. La discussion du système est alors élémentaire : il n'est compatible que si

$$b'_{r+1} = \cdots = b'_p = 0$$

(ce sont les $p - r$ **conditions de compatibilité** *cf.* théorème **4-5**.1). Si ces conditions sont remplies, on peut choisir les r premières inconnues comme inconnues principales (attention ! elles ont sans doute été renumérotées...) et on s'aperçoit que, pour un choix arbitraire des inconnues secondaires $x_{r+1}, \ldots x_n$, les valeurs de $x_1, \ldots, x_r$ s'obtiennent en résolvant un système triangulaire : les $(x_i)_{1 \leqslant i \leqslant r}$ s'obtiennent de manière unique comme fonctions affines des inconnues secondaires.

Terminons cette section avec quelques remarques "numériques", en supposant $\mathbb{K} = \mathbb{R}$ (ou $\mathbb{C}$) et, pour simplifier $r = p = n$ (système de Cramer). Lorsqu'on débute les manipulations élémentaires sur les lignes de la matrice A', dont les coefficients sont entachés d'erreurs (résultats de mesures, de calculs approchés etc...), en supposant par exemple $a_{1,1} \neq 0$, on effectue

$$L_i \longleftarrow L_i - \frac{a_{i,1}}{a_{1,1}} L_1 \quad \text{pour } 2 \leqslant i \leqslant n$$

(on dit que l'on a choisi le coefficient diagonal $a_{1,1}$ comme "pivot"). Une erreur ε sur un coefficient de L_1 sera multipliée par le coefficient $\dfrac{a_{i,1}}{a_{1,1}}$, ce qui peut être très mauvais si ce coefficient est très grand (d'autant plus qu'on n'en est qu'à la première étape du processus de triangulation du système !). Deux stratégies sont possibles pour être certain que ce coefficient reste inférieur à 1 en valeur absolue (ou en module) :

- Stratégie du pivot partiel : on cherche dans la première colonne de A' un coefficient $a_{i,1}$ **maximal en valeur absolue** (ou module) et on commence par effectuer

  $$L_1 \leftrightarrow L_i$$

  C'est donc le coefficient $a_{i,1}$ qui est choisi comme "pivot". On opèrera de la même façon à chaque étape du processus (avec les notations de la démonstration du théorème **4-6**.2, on cherchera un coefficient de module maximal dans la première colonne de A_1 etc...)

- Stratégie du pivot total : on gagne en précision en cherchant parmi les coefficients de la matrice du système A un coefficient $a_{i,j}$ maximal en module et on commencera par effectuer sur la matrice complète A' les manipulations

$$C_1 \leftrightarrow C_j \text{ (attention ! renumérotation des inconnues) puis } L_1 \leftrightarrow L_i$$

On recommencera ensuite avec la matrice A_1... Il est essentiel en vue d'une programmation de conserver une trace des permutations de variables. On peut par exemple initialiser un vecteur à $(1, 2, \ldots, n)$ et opérer sur les composantes de ce vecteur les permutations effectuées sur les colonnes de A.

Un exemple montrer l'importance du choix du pivot. Imaginons un calculateur rudimentaire opérant en virgule flottante avec 3 chiffres significatifs. Pour notre machine on a

$$1 + 10^{-4} = 1$$

Le système d'équations

$$\begin{cases} 10^{-4}x_1 & + & x_2 & = & 1 \\ x_1 & + & x_2 & = & 2 \end{cases}$$

a pour solutions $x_1 = 1,00010\ldots$ et $x_2 = 0,9990\ldots$, indistinguables pour notre machine de $x_1 = x_2 = 1$. On vérifie aisément que c'est bien cette solution que donnerait la méthode du pivot avec le choix $a_{1,1} = 1$. Le choix $a_{1,1} = 10^{-4}$ serait désastreux, puisqu'il amènerait au système équivalent

$$\begin{cases} 10^{-4}x_1 & + & x_2 & = & 1 \\ & & (1 - 10^4)x_2 & = & 2 - 10^4 \end{cases}$$

Pour notre machine, la seconde équation donnerait $x_2 = 1$. En reportant aveuglément dans la première égalité, on obtient $x_1 = 0$!

4-6.4.2 Calcul du déterminant d'une matrice carrée

Pour les matrices carrées, le résultat du corollaire **4-6**.2 donne, avec les notations de cet énoncé

$$\det A = \pm \prod_{i=1}^{n} \alpha_{i,i}$$

où le signe $\pm$ dépend de la parité du nombre transpositions de colonnes (ou de lignes en cas de stratégie du pivot total) effectuées pour amener la matrice A à une forme triangulaire. On gardera une trace de ces permutations en initialisant une variable signe à $+1$ et en changeant le contenu en son opposé à chaque transposition de ligne ou colonne effectuée sur la matrice. Il est à noter que le nombre d'opérations de base (addition ou multiplication) effectuées pour obtenir la valeur du déterminant d'une matrice de type (n, n) est en $O(n^3)$, à rapprocher de la quantité astronomique $n!$ que donnerait une stratégie aveugle de développement par rapport à une ligne quelconque.

4-6.4.3 Calcul de l'inverse d'une matrice carrée

Si $A \in \mathcal{GL}_n(\mathbb{C})$, mettons en oeuvre la stratégie du pivot partiel. On obtient une matrice triangulaire supérieure T et une matrice carrée M, produit de matrices élémentaires

$$M = P_k P_{k-1} \cdots P_2 P_1$$

telles que

$$MA = T$$

Pour obtenir la matrice M, il suffit de remarquer que

$$M = P_k P_{k-1} \cdots P_2 P_1 I_n$$

est le résultat de la suite de manipulations élémentaires que l'on a effectuées sur A, opérées en partant de la matrice identité. On réévalue cette matrice à chaque étape du processus du pivot. On obtient alors évidemment

$$TA^{-1} = M$$

ce qui permet de remplir la matrice A^{-1} ligne par ligne, en partant de la dernière, en résolvant le système (triangulaire)

$$T \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} = M = \begin{pmatrix} M_1 \\ \vdots \\ M_n \end{pmatrix}$$

en notant L_i (resp. M_i) la $i^{\text{ème}}$ ligne de A^{-1} (resp. de M).

Dans le cas de la mise en oeuvre d'une stratégie du pivot total, on obtient une égalité

$$MAP_\sigma = T$$

où P_σ est une matrice de permutation. On aura cette fois

$$T(AP_\sigma)^{-1} = M$$

ce qui permet d'obtenir les lignes de $D = (AP_\sigma)^{-1}$, puis de calculer la matrice $A^{-1} = P_\sigma (AP_\sigma)^{-1}$, en opérant des permutations sur les lignes de D, pour peu qu'on ait pris le soin de garder une trace de σ (ou de σ^{-1}).

4-6.4.4 Le rang caractérise les matrices à équivalence près

On obtient ici une version "algorithmique" du résultat du théorème **4-2.9**. Si A est une matrice $(a_{ij})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K})$, de rang $r > 0$, le pivot de Gauss montre que A est équivalente à une matrice

$$B = (b_{ij})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}} = \left(\begin{array}{ccc|c} b_{1,1} & ? & ? & \\ 0 & \ddots & ? & ? \\ 0 & 0 & b_{r,r} & \\ \hline 0 & \cdots & 0 & 0_{p-r,n-r} \end{array} \right)$$

Il est alors très simple de transformer (par manipulations élémentaires, donc par équivalence) cette matrice en $J_r^{n,p} = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{p-r,r} & 0_{p-r,n-r} \end{pmatrix}$. Il suffit d'opérer par exemple

$$C_1 \leftarrow \frac{1}{b_{1,1}} C_1 \quad \text{puis} \quad C_i \leftarrow C_i - b_{1,i} C_1 \text{ pour } i = 2 \cdots n$$

et d'opérer ensuite de même pour les colonnes 2 à r. Remarquons que c'est uniquement à ce stade que nous avons fait intervenir des manipulations élémentaires de type $\mathcal{L}_2$ ou $\mathcal{C}_2$ (donc des matrices élémentaires de type $\mathcal{T}_2$). Nous verrons, dans le paragraphe suivant et dans le cadre des matrices carrées, que ce qui oblige à faire intervenir les matrices $\mathcal{T}_2$, c'est l'existence du déterminant.

4-6.5 Exercice : générateurs de $\mathcal{GL}_n(\mathbb{K})$

Nous commencerons par une interprétation géométrique des matrices carrées de type $\mathcal{T}_1$ ou $\mathcal{T}_2$.

4-6.5.1 Transvections, dilatations

Soit $\mathbb{E}$ un $\mathbb{K}$-ev de dimension n, et $\mathbb{H}$ un hyperplan de $\mathbb{E}$. On s'intéresse aux automorphismes u de $\mathbb{E}$ qui laissent invariants tous les éléments de $\mathbb{H}$. Si on laisse de côté l'identité, montrer qu'il existe un vecteur $a \neq 0_{\mathbb{E}}$ et une forme linéaire non nulle φ tels que

$$\forall x \in \mathbb{E} \quad u(x) = x + \varphi(x) a$$

On distingue alors deux cas :

- Si $a \in \mathbb{H}$, on dit que u est une **transvection d'hyperplan** $\mathbb{H}$. Montrer qu'alors il existe une base de $\mathbb{E}$ dans laquelle la matrice de u s'écrit $M_{i,j}(\lambda)$, avec $i \neq j$ quelconques et $\lambda \neq 0$ arbitraire. C'est pour cela que les matrices élémentaires de type $\mathcal{T}_1$ sont appelées matrices élémentaires de transvection.

- Si $a \notin \mathbb{H}$, montrer que $\alpha = \varphi(a) + 1 \neq 0$ et qu'il existe une base de $\mathbb{E}$ où la matrice de u s'écrit $D_i(\alpha)$ avec i arbitraire. L'automorphisme u est alors appelé **dilatation de base l'hyperplan** $\mathbb{H}$, **de direction** $\text{vect}(a)$ **et de rapport** α. Il est obtenu en recollant l'identité sur $\mathbb{H}$ et l'homothétie de rapport α sur la direction supplémentaire $\text{vect}(a)$. Les matrices élémentaires de type $\mathcal{T}_2$ sont aussi appelées matrices élémentaires de dilatation.

4-6.5.2 Lemme du pivot de Jordan

On note $\mathbb{G}$ l'ensemble des matrices de $\mathcal{GL}_n(\mathbb{K})$ pouvant s'écrire comme produits de matrices élémentaires de type $\mathcal{T}_1$. Montrer que $\mathbb{G}$ est un sous-groupe de $\mathcal{GL}_n(\mathbb{K})$ et que $\mathbb{G} \subset \mathcal{SL}_n(\mathbb{K})$ (*cf.* section **4-4.3**)

LEMME 4-6.3 On suppose $A \in \mathcal{GL}_n(\mathbb{K})$ avec $n \geqslant 2$. Il existe alors deux matrices G_1 et G_2 dans $\mathbb{G}$ telles que

$$G_1AG_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

Démonstration : On notera qu'il s'agit de transformer la matrice A uniquement par des manipulations de type $\mathcal{L}_1$ ou $\mathcal{C}_1$. S'il existe un indice $i > 1$ avec $a_{i,1} \neq 0$, on choisit

$$\lambda = \frac{1 - a_{1,1}}{a_{i,1}}$$

et on effectue les manipulations

$$L_1 \leftarrow L_1 - \lambda L_i$$

$$\text{puis} \quad L_j \leftarrow L_j - a_{j,1}L_1 \text{ pour } j = 2 \cdots n$$

$$\text{et} \quad C_j \leftarrow C_j - a_{1,j}C_1 \text{ pour } j = 2 \cdots n$$

Si tous les coefficients $a_{i,1}$ sont nuls pour $i > 1$, il existe nécessairement des indices i et $j \geqslant 2$ avec $a_{i,j} \neq 0$, car sinon A ne serait pas inversible. Il suffit alors d'opérer au préalable

$$C_1 \leftarrow C_1 + C_j$$

pour être ramené au cas précédent. ■

4-6.5.3 Générateurs de $\mathcal{GL}_n(\mathbb{K})$

Déduire de ce qui précède que, si $A \in \mathcal{GL}_n(\mathbb{K})$, il existe des matrices G_1 et G_2 de $\mathbb{G}$ telles

$$G_1AG_2 = \begin{pmatrix} I_{n-1} & 0 \\ 0 & \det A \end{pmatrix}$$

En déduire que $\mathbb{G} = \mathcal{SL}_n(\mathbb{K})$ et que toute matrice de $\mathcal{GL}_n(\mathbb{K})$ peut s'écrire comme produit de matrices élémentaires de transvections et d'une matrice de dilatation. La version géométrique de ce théorème est l'existence d'une écriture comme produit de transvections et de dilatations pour tout automorphisme d'un espace vectoriel de dimension finie.

EXERCICE 4-6.4 Soit $\phi : \mathcal{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ un morphisme de groupes multiplicatifs tel que $\phi\left((a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}\right)$ s'exprime comme polynôme des n^2 indéterminées $a_{i,j}$. Montrer qu'il existe un entier p tel que

$$\forall M \in \mathcal{GL}_n(\mathbb{R}) \quad \phi(M) = (\det M)^p$$

4-7 Exercices

EXERCICE 4-7.1 Soient A et B dans $\mathcal{M}_n(K)$. On suppose qu'il existe $p \in \mathbb{N}$ tel que $A^pB + A + B = 0$. Montrer que A et B commutent.

EXERCICE 4-7.2 Soit p un nombre premier et F_p un corps à p éléments. Déterminer le cardinal de $\mathcal{GL}_n(F_p)$.

EXERCICE 4-7.3 Etudier le rang et inverser le cas échéant la matrice complexe

$$A = \begin{pmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 1 & 1+a_n \end{pmatrix}$$

EXERCICE 4-7.4 Déterminer le rang, le déterminant et la trace de l'application de $\mathcal{M}_n(K)$ dans $\mathcal{M}_n(K)$ $M \mapsto AMB$, où $A, B \in \mathcal{M}_n(K)$ sont fixées. Même question avec $M \mapsto {}^tM$.

EXERCICE 4-7.5 Calculer les déterminants des matrices :

$$\begin{pmatrix} a_1+b_1 & b_1 & \cdots & b_1 \\ b_2 & a_2+b_2 & \cdots & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ b_n & \cdots & b_n & a_n+b_n \end{pmatrix} \quad \begin{pmatrix} 1 & x & x^2 & \cdots & x^n \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & n+1 \\ \vdots & & & & \vdots \\ 1 & 2^{n-1} & 3^{n-1} & \cdots & (n+1)^{n-1} \end{pmatrix}$$

$$\begin{pmatrix} 1+x_1 & \cdots & 1+x_n \\ 1+x_1^2 & \cdots & 1+x_n^2 \\ \vdots & & \vdots \\ 1+x_1^n & \cdots & 1+x_n^n \end{pmatrix}.$$

EXERCICE 4-7.6 Soient A et B dans $\mathcal{M}_n(\mathbb{Q})$. Montrer l'équivalence des propriétés :
i) $\exists\, P \in \mathcal{GL}_n(\mathbb{C}) \quad B = P^{-1}AP$.
ii) $\exists\, P \in \mathcal{GL}_n(\mathbb{R}) \quad B = P^{-1}AP$.
iii) $\exists\, P \in \mathcal{GL}_n(\mathbb{Q}) \quad B = P^{-1}AP$.

EXERCICE 4-7.7 Pour $x \in \mathbb{R}$ et $n \in \mathbb{N}^*$ on pose

$$D_n(x) =: \begin{vmatrix} 1+x^2 & x & 0 & \cdots & 0 \\ x & 1+x^2 & x & 0 & \vdots \\ 0 & x & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & x \\ 0 & \cdots & \cdots & x & 1+x^2 \end{vmatrix}$$

Montrer que $D_n(x)$ est un polynôme en x. Quel est son degré ? Trouver une relation de récurrence entre les D_n. Calculer les D_n.

EXERCICE 4-7.8 Pour $n \geqslant 1$, on note b_n le nombre de permutations de $\{1, \cdots, n\}$ sans point fixe (on posera $b_0 = 1$). Ecrire $n!$ comme combinaison linéaire des b_k. Soit la matrice : $P = \left(C_i^j\right)_{0 \leqslant i,j \leqslant n}$. Expliciter P^{-1}. Déterminer la limite de $b_n/n!$.

EXERCICE 4-7.9 Soient $A, B, C, D \in \mathcal{M}_n(\mathbb{R})$ avec $C^tD + D^tC = 0$.

1. Montrer que, si D est inversible, $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A^tD + B^tC)$.

2. Montrer que ce résultat n'est pas valable en général si D n'est pas inversible, mais que l'égalité élevée au carré est toujours vérifiée.

EXERCICE 4-7.10 Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{n,m}(\mathbb{R})$ et $C = AB$. Que peut on dire de $\det C$ si $n < m$? Montrer que, pour $n \geqslant m$:

$$\det C = \sum_{r_1 < r_2 < \cdots < r_m} \det A \begin{pmatrix} 1 & 2 & \cdots & m \\ r_1 & r_2 & \cdots & r_m \end{pmatrix} \times \det B \begin{pmatrix} r_1 & r_2 & \cdots & r_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

où $M \begin{pmatrix} i & \cdots & j \\ k & \cdots & l \end{pmatrix}$ est la matrice extraite de M à l'aide des lignes $i \cdots j$ et des colonnes $k \cdots l$. Qu'obtient-on pour $B = {}^tA$?

EXERCICE 4-7.11 Soient A et $B \in \mathcal{M}_n(\mathbb{R})$ et soit $M = \begin{pmatrix} A & B & B & \cdots & B \\ B & A & B & \cdots & B \\ \vdots & & \ddots & \cdots & \vdots \\ B & \cdots & B & A & B \\ B & \cdots & B & B & A \end{pmatrix}$ appartenant à $\mathcal{M}_{np}(\mathbb{R})$.
Soit $E = \{\lambda \in \mathbb{R} \mid \det(A - \lambda B) \neq 0\}$. Trouver une condition nécessaire et suffisante sur E pour que M soit inversible et calculer alors M^{-1}.

EXERCICE 4-7.12 Factorisation LU : Soit $M \in \mathcal{M}_n(K)$. On appelle mineurs principaux de M les déterminants $(\Delta_i)_{1 \leqslant i \leqslant n}$ obtenus en rayant les $n - i$ dernières lignes et colonnes de M. Montrer que, si M est inversible, il y a équivalence entre les propositions :
i) Les mineurs principaux de M sont tous non nuls.
ii) M peut se factoriser sous la forme $M = LU$ où L est triangulaire inférieure à diagonale formée de 1 et U est triangulaire supérieure.
Montrer qu'une telle décomposition est alors unique.

EXERCICE 4-7.13 Soit E l'espace vectoriel $\mathbb{C}^{\mathbb{R}}$. Montrer que n fonctions $f_1, \cdots, f_n$ de E sont linéairement indépendantes ssi $\exists (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ avec $\det(f_i(x_j)) \neq 0$.

EXERCICE 4-7.14 Si $A \in \mathcal{M}_n(K)$ on note $c(A)$ sa comatrice.

1. Comparer les rangs de A et $c(A)$.

2. Comparer les déterminants de A et $c(A)$.

3. Pour $A, B \in \mathcal{M}_n(K)$ comparer $c(AB)$ et $c(A)c(B)$.

4. Si A est non inversible et $c(A) = (A_{i,j})$ montrer que $\forall i, j, k, l \;\; A_{i,j}A_{k,l} = A_{k,j}A_{i,l}$.

Chapitre 5

Polynômes, réduction des endomorphismes

5-1 Arithmétique des polynômes

Le programme de première année étudie l'arithmétique des anneaux $(\mathbb{Z},+,\times)$ et $(\mathbb{K}[X],+,\times)$. L'analogie entre les deux théories s'explique par l'existence d'une division euclidienne dans chacun de ces anneaux. Avec la théorie des idéaux, nous repasserons en revue les principaux résultats concernant la divisibilité dans l'anneau $\mathbb{K}[X]$.

5-1.1 Idéal d'un anneau commutatif

5-1.1.1 Anneau, sous-anneau

Soit $(\mathbb{A},+,\times)$ un anneau. Rappelons que $(\mathbb{A},+)$ est un groupe commutatif, dont l'élément neutre est noté $0_{\mathbb{A}}$, que la multiplication est associative, distributive par rapport à l'addition et possède un élément neutre, appelé élément unité de l'anneau $\mathbb{A}$, que nous noterons $1_{\mathbb{A}}$, avec $1_{\mathbb{A}} \neq 0_{\mathbb{A}}$. Lorsque la multiplication est commutative, on parlera d'anneau commutatif.

DÉFINITION 5-1.1 (sous-anneau) *Une partie $\mathbb{B} \subset \mathbb{A}$ est un sous-anneau de $\mathbb{A}$ si et seulement si $(\mathbb{B},+)$ est un sous-groupe de $(\mathbb{A},+)$ (donc $\mathbb{B}$ est stable pour l'addition et le passage à l'opposé), si $1_{\mathbb{A}} \in \mathbb{B}$ et si $\mathbb{B}$ est stable pour la multiplication.*

Il est alors clair que $(\mathbb{B},+,\times)$ est également une structure d'anneau. On vérifie qu'une intersection d'une famille de sous-anneaux de $\mathbb{A}$ est encore un sous-anneau de $\mathbb{A}$.

DÉFINITION 5-1.2 (morphisme d'anneaux) *Si $(\mathbb{A}_1,+,\times)$ et $(\mathbb{A}_2,+,\times)$ sont deux anneaux, une application*

$$\varphi : \mathbb{A}_1 \to \mathbb{A}_2$$

est un morphisme d'anneaux ssi φ est un morphisme pour les lois additives et multiplicatives et si

$$\varphi(1_{\mathbb{A}_1}) = 1_{\mathbb{A}_2}$$

Vérifier qu'une composée de morphismes d'anneaux est encore un morphisme et qu'images directes ou réciproques de sous-anneaux sont encore des sous-anneaux. Il en est en particulier ainsi pour l'image d'un morphisme d'anneaux.

Sur un produit d'une famille d'anneaux est définie une structure "naturelle" d'anneau, pour laquelle les opérations sont définies coordonnée par coordonnée.

DÉFINITION 5-1.3 (anneau intègre) *Un élément non nul $a \in \mathbb{A}$ est dit diviseur de zéro à gauche s'il existe $b \in \mathbb{A}$ non nul avec $ab = 0_{\mathbb{A}}$ (b est alors évidemment diviseur de zéro à droite, la distinction n'ayant pas de sens si l'anneau est commutatif). Comme*

$$\forall x, y \in \mathbb{A} \quad ax = ay \Leftrightarrow a(x-y) = 0_{\mathbb{A}}$$

dire que a n'est pas diviseur de zéro à gauche équivaut à dire que a est simplifiable (ou régulier) à gauche (pour la multiplication, évidemment !).

Un anneau est dit intègre s'il ne possède pas de diviseurs de zéros.

5-1.1.2 Groupe des unités d'un anneau

DÉFINITION 5-1.4 *Un élément a d'un anneau $(\mathbb{A},+,\times)$ est dit unité de $\mathbb{A}$ ssi a est inversible (à droite et à gauche) pour la multiplication.*

Attention au vocabulaire ! Une unité d'un anneau n'est pas forcément l'élément unité ! Par exemple , dans $\mathbb{Z}$ les unités sont 1 et -1. Dans $\mathbb{K}[X]$, des considérations de degré montrent que les unités sont exactement les polynômes de degré zéro, c'est-à-dire constants non nuls.

La vérification du théorème suivant est immédiate :

THÉORÈME 5-1.5 L'ensemble $U(\mathbb{A})$ des unités d'un anneau $(\mathbb{A},+,\times)$ est un groupe pour la multiplication, appelé groupe des unités de $\mathbb{A}$.

Si $\mathbb{A}$ est un corps, on a évidemment $U(\mathbb{A}) = \mathbb{A}^* = \mathbb{A} - \{0_{\mathbb{A}}\}$.

5-1.1.3 Divisibilité dans un anneau commutatif intègre

Nous nous placerons désormais dans un anneau $(\mathbb{A}, +, \times)$ commutatif sans diviseur de zéro, ce qui simplifie considérablement les règles de calcul, notamment pour les simplifications.

DÉFINITION 5-1.6 *Si a et $b \in \mathbb{A}$, on dit que b divise a si et seulement s'il existe $c \in \mathbb{A}$ avec $a = bc$.*

On note en abrégé $b|a$, qu'on lit indifféremment "b divise a", "b est un diviseur de a" ou "a est un multiple de b". On notera cependant une ambiguïté si $a = 0_{\mathbb{A}}$: tout élément $b \in \mathbb{A}$ divise $0_{\mathbb{A}}$ puisqu'on a toujours

$$0_{\mathbb{A}} = 0_{\mathbb{A}}b$$

mais on notera que "être un diviseur de zéro" a un autre sens pour nous. Ceci n'est pas bien grave, puisque la relation de divisibilité n'est intéressante que sur les éléments non nuls de l'anneau.

La relation de divisibilité est évidemment réflexive et transitive. Elle n'est pas véritablement antisymétrique (bien que l'on sache par exemple dans le cas de $\mathbb{Z}$ qu'en restriction à $\mathbb{N}^*$ ce soit une relation d'ordre) :

THÉORÈME 5-1.7 Si $(\mathbb{A}, +, \times)$ est un anneau intègre, pour a et $b \in \mathbb{A}$

$$a|b \text{ et } b|a \Leftrightarrow \exists\, u \in U(\mathbb{A}) \quad b = ua$$

Démonstration : Si $b = ua$, on a évidemment $a = u^{-1}b$, et donc $a|b$ et $b|a$. Réciproquement, s'il existe c et $d \in \mathbb{A}$ avec $a = cb$ et $b = da$, on obtient

$$a = cda \Rightarrow a(1_{\mathbb{A}} - cd) = 0_{\mathbb{A}}$$

Comme l'anneau est intègre, on en déduit $a = 0_{\mathbb{A}}$ (donc aussi $a = b$) ou $cd = 1_{\mathbb{A}}$. Comme a et b jouent des rôles symétriques on aurait aussi $dc = 1_{\mathbb{A}}$ ce qui montre que c et d sont deux unités de $\mathbb{A}$. ■

On dit alors que a et b **sont deux éléments associés de l'anneau $\mathbb{A}$**. On définit ainsi une relation d'équivalence sur l'anneau $\mathbb{A}$. Dans le cas de $\mathbb{Z}$, la classe d'équivalence de a est $\{\pm a\}$ et possède donc un unique représentant dans $\mathbb{N}$, égal à $|a|$. Dans $\mathbb{K}[X]$, la classe d'un polynôme non nul P est $\{\alpha P, \quad \alpha \in \mathbb{K}^*\}$ qui possède un unique représentant de coefficient dominant égal à 1 (représentant "normalisé", on dit parfois "unitaire"). Vis-à-vis de la relation de divisibilité, on ne peut distinguer deux éléments associés d'un anneau, ils sont "équivalents". Les unités d'un anneau divisent tout élément de cet anneau. En particulier, dans un corps, la relation de divisibilité est triviale : tout élément non nul divise tout élément ! C'est pour cela que lorsqu'on a, par exemple, à résoudre des problèmes de divisibilité dans $\mathbb{Z}$, il est fortement conseillé de ne faire intervenir que des entiers relatifs, mais pas de rationnels : on risque sinon de se laisser entraîner par ses notations et de croire qu'une égalité $a = bcd$ avec $a, b \in \mathbb{Z}$ et $c, d \in \mathbb{Q}$ peut permettre de conclure que b divise a dans $\mathbb{Z}$.

DÉFINITION 5-1.8 *Un élément a d'un anneau commutatif intègre $\mathbb{A}$ est dit* ***irréductible*** *si a n'est pas une unité de $\mathbb{A}$ et si tout diviseur de a est soit une unité, soit un élément associé à a (on dit aussi que a n'a que des diviseurs "triviaux").*

Dans $(\mathbb{Z},+,\times)$, les irréductibles sont les entiers naturels premiers et leurs opposés. Dans $\mathbb{K}[X]$, la situation dépend du corps $\mathbb{K}$:

Des considérations de degré montrent qu'un polynôme de degré 1 est toujours irréductible. Si $\mathbb{K}=\mathbb{C}$ la réciproque est vraie. Ceci est conséquence du théorème de D'Alembert stipulant que tout polynôme de $\mathbb{C}[X]$ de degré >0 possède au moins une racine, et parce qu'on a l'équivalence, valable sur un corps quelconque et conséquence de l'existence d'une division euclidienne

$$\forall P\in\mathbb{K}[X]\quad \forall a\in\mathbb{K}\quad (X-a)|P \Leftrightarrow P(a)=0$$

Dans le cas où $\mathbb{K}=\mathbb{R}$, comme pour $P\in\mathbb{R}[X]$ et $z\in\mathbb{C}-\mathbb{R}$ on a

$$P(z)=0\Leftrightarrow P(\overline{z})=0$$

et $(X-z)(X-\overline{z})\in\mathbb{R}[X]$, on montre aisément que les polynômes irréductibles sont les polynômes de degré 1 et les polynômes de degré 2 à discriminant strictement négatif.

Dans $\mathbb{Q}[X]$, il existe des polynômes irréductibles de degré aussi élevé qu'on le souhaite. On peut par exemple montrer (voir exercice **5-8**.4) que, pour p entier premier, le polynôme

$$1+X+X^2+\cdots+X^{p-1}$$

est irréductible dans $\mathbb{Q}[X]$.

5-1.1.4 Idéal d'un anneau commutatif

Si $(\mathbb{A}_1,+,\times)$ et $(\mathbb{A}_2,+,\times)$ sont deux anneaux (non nécessairement commutatifs), et $\varphi:\mathbb{A}_1\to\mathbb{A}_2$ un morphisme d'anneaux, son noyau

$$\ker\varphi=\varphi^{-1}(0_{\mathbb{A}_2})$$

est un sous-groupe de $(\mathbb{A}_1,+)$ (comme tout noyau d'un morphisme de groupes !), qui possède une propriété plus forte que la stabilité pour la multiplication (*cf.* section **2-2.3**) :

$$\forall a\in\mathbb{A}\quad \forall x\in\ker\varphi\quad ax\in\ker\varphi\quad \text{et}\quad xa\in\ker\varphi$$

Ceci amène la définition :

DÉFINITION 5-1.9 (idéal d'un anneau commutatif) *Si $(\mathbb{A},+,\times)$ est un anneau commutatif*[1], *une partie $\mathcal{I}\subset\mathbb{A}$ est un idéal de $\mathbb{A}$ si et seulement si $(\mathcal{I},+)$ est un sous-groupe de $(\mathbb{A},+)$ et si*

$$\forall a\in\mathbb{A}\quad \forall x\in\mathcal{I}\quad ax\in\mathcal{I}$$

[1]Si l'anneau n'est pas commutatif, un sous-groupe additif $\mathcal{I}$ est dit idéal à **gauche** si

$$\forall a\in\mathbb{A}\quad \forall x\in\mathcal{I}\quad ax\in\mathcal{I}$$

On définit de même la notion d'idéal à droite. Un idéal à droite et à gauche est dit **bilatère.** Le noyau d'un morphisme d'anneaux est toujours un idéal bilatère de l'anneau de départ

REMARQUE 5-1.10 $\{0_{\mathbb{A}}\}$ et $\mathbb{A}$ sont toujours des idéaux de $\mathbb{A}$. Pour qu'un idéal $\mathcal{I}$ de $\mathbb{A}$ soit égal à $\mathbb{A}$ il suffit d'ailleurs que $\mathcal{I}$ contienne $1_{\mathbb{A}}$, ou une unité de $\mathbb{A}$, puisque si $u \in U(\mathbb{A}) \cap \mathcal{I}$

$$\forall a \in \mathbb{A} \quad a = (au^{-1})\, u \in \mathcal{I}$$

En particulier, un corps commutatif $\mathbb{K}$ ne possède que deux idéaux $\{0_{\mathbb{K}}\}$ et $\mathbb{K}$. Il en résulte, d'après le résultat évoqué en début de cette section, qu'un morphisme de corps est toujours injectif.

REMARQUE 5-1.11 Une partie $\mathcal{I}$ non vide de $\mathbb{A}$ est un idéal ssi $\mathcal{I}$ est stable pour l'addition et vérifie $\forall a \in \mathbb{A} \quad \forall x \in \mathcal{I} \quad ax \in \mathcal{I}$. En effet, la propriété de stabilité pour le passage à l'opposé est conséquence de l'égalité $-x = (-1_{\mathbb{A}})\, x$.

Comme dans le cadre des sous-espaces vectoriels, on démontre aisément les résultats suivants :

THÉORÈME 5-1.12 L'intersection d'une famille d'idéaux d'un anneau commutatif est un idéal. Ceci permet de définir la notion d'idéal engendré par une famille d'éléments (ou une partie) de l'anneau.

THÉORÈME 5-1.13 Si $a \in \mathbb{A}$, l'idéal engendré par $\{a\}$ est l'ensemble des multiples de a dans $\mathbb{A}$. On le note indifféremment (a) ou $a\mathbb{A}$ (ou $\mathbb{A}a$ puisque l'anneau est supposé commutatif) :

$$(a) = a\mathbb{A} = \{ax,\ x \in \mathbb{A}\}$$

DÉFINITION 5-1.14 *Un idéal engendré par un élément est dit* ***principal*** *(ou* ***monogène****). L'idéal principal engendré par a est donc l'ensemble des multiples de a.*

EXERCICE 5-1.15 Montrer de même que l'idéal engendré par une famille finie $(a_i)_{1 \leqslant i \leqslant p}$ d'éléments de $\mathbb{A}$ est

$$(a_1, a_2, \ldots, a_p) = \{x \in \mathbb{A} \mid \exists x_1, \ldots x_p \in \mathbb{A} \quad x = a_1 x_1 + \cdots + a_p x_p\}$$

EXERCICE 5-1.16 Si $\mathcal{I}_1$ et $\mathcal{I}_2$ sont deux idéaux de $\mathbb{A}$, on appelle somme de ces idéaux l'idéal (noté $\mathcal{I}_1 + \mathcal{I}_2$) engendré par la réunion $\mathcal{I}_1 \cup \mathcal{I}_2$. Caractériser les éléments de cette somme.

EXERCICE 5-1.17 Si $\mathcal{I}$ est un idéal de l'anneau **commutatif** $\mathbb{A}$, on appelle **radical** de $\mathcal{I}$ l'ensemble

$$\sqrt{\mathcal{I}} = \{x \in \mathbb{A} \mid \exists n \in \mathbb{N}^* \quad x^n \in \mathcal{I}\}$$

Montrer que cet ensemble est un idéal de $\mathbb{A}$. En particulier, l'ensemble des éléments nilpotents de $\mathbb{A}$ (éléments dont une puissance est nulle) est un idéal.

La notion de divisibilité est liée à la notion d'idéal (principal) :

THÉORÈME 5-1.18 Dans un anneau $\mathbb{A}$ commutatif, pour a et $b \in \mathbb{A}$, on a

$$b | a \Leftrightarrow (a) \subset (b)$$

Si de plus $\mathbb{A}$ est intègre, a et b sont associés ssi ils engendrent le même idéal.

Démonstration : Si $(a) \subset (b)$, en particulier $a \in (b)$ donc a est un multiple de b. Réciproquement, si b est un diviseur de a, il est clair que tout multiple de a est aussi multiple de b. ■

5-1.1.5 Idéaux de $\mathbb{Z}$ et de $\mathbb{K}[X]$

L'existence d'un division euclidienne dans chacun de ces anneaux a une conséquence très importante, qui va permettre d'unifier la théorie du PGCD (et du PPCM) dans ces deux anneaux :

THÉORÈME 5-1.19 Tout idéal de $\mathbb{Z}$ ou de $\mathbb{K}[X]$ est engendré par un élément.

Démonstration : En particulier, puisque deux éléments qui engendrent le même idéal sont associés, tout idéal de $\mathbb{K}[X]$ non réduit à $\{0\}$ possède un unique générateur normalisé. De même, un idéal de $\mathbb{Z}$ possède un unique générateur entier naturel.

Faisons la démonstration dans $\mathbb{K}[X]$, elle est analogue dans $\mathbb{Z}$, pour peu que l'on remplace le degré par la valeur absolue. Soit $\mathcal{I}$ un idéal de $\mathbb{K}[X]$. Si $\mathcal{I} = \{0\}$, on a aussi $\mathcal{I} = (0)$. Sinon l'ensemble des degrés des polynômes non nuls de $\mathcal{I}$ est une partie non vide de $\mathbb{N}$, qui possède un plus petit élément $d \geqslant 0$. Si $A \in \mathcal{I}$ est choisi de degré d, on a évidemment $(A) \subset \mathcal{I}$. Si $P \in \mathcal{I}$, effectuons la division euclidienne de P par A :

$$\exists Q, R \in \mathbb{K}[X] \quad P = AQ + R \quad \text{et} \quad R = 0 \text{ ou } d^\circ R < d^\circ A$$

Comme $\mathcal{I}$ est un idéal, $R = P - AQ \in \mathcal{I}$, ce qui impose, vu le choix de A (degré minimal), $R = 0$. En conséquence $P \in (A)$ et $\mathcal{I} \subset (A)$. ■

Un anneau commutatif intègre où tout idéal est principal est appelé **anneau principal**. Nous verrons dans la section suivante que l'on peut mener dans un anneau principal une théorie du PGCD comme dans $\mathbb{Z}$ ou $\mathbb{K}[X]$, en utilisant uniquement la notion d'idéal.

5-1.2 Arithmétique dans $\mathbb{Z}$ et $\mathbb{K}[X]$

Nous nous placerons ici dans un anneau principal $(\mathbb{A}, +, \times)$ quelconque. Dans la pratique, ce sera le plus souvent $\mathbb{Z}$ ou $\mathbb{K}[X]$.

5-1.2.1 Plus grand diviseur commun

Nous étudierons d'abord la notion de PGCD pour deux éléments de $\mathbb{A}$.

THÉORÈME 5-1.20 (et définition) Si a et $b \in \mathbb{A}$, l'ensemble

$$(a) + (b) = \{ax + by \mid x, y \in \mathbb{A}\}$$

est un idéal de $\mathbb{A}$. Un générateur quelconque de cet idéal est appelé un PGCD de a et b. Si

$$(d) = (a) + (b)$$

on notera

$$d = \text{PGCD}\,(a,b) \quad \textbf{ou} \quad d = a \wedge b$$

Démonstration : Il est à noter que l'égalité $d = \text{PGCD}\,(a,b)$ n'en est pas vraiment une, puisqu'un idéal de $\mathbb{A}$ n'a pas en général un unique générateur. En particulier, de deux écritures $d_1 = \text{PGCD}\,(a,b)$ et $d_2 = \text{PGCD}\,(a,b)$, on se gardera bien de déduire $d_1 = d_2$ mais on traduira "d_1 et d_2 sont associés". Bien sûr, dans le cas de $\mathbb{Z}$, si on décide qu'un PGCD est toujours positif, il n'y a plus d'ambiguïté. De même dans $\mathbb{K}[X]$ si on choisit systématiquement dans un idéal non nul son unique générateur normalisé.

Vérifier que $(a) + (b)$ est un idéal est immédiat. ∎

Reste à comprendre la terminologie "PGCD". Cela est dû à la caractérisation suivante du PGCD, et à la remarque qui en découle :

PROPOSITION 5-1.21 Pour a, b et $d \in (\mathbb{A}, +, \times)$, on a

$$d = \text{PGCD}\,(a,b) \Leftrightarrow \begin{cases} d \textbf{ divise } a \textbf{ et } b \\ \exists\, u, v \in \mathbb{A} \quad d = au + bv \end{cases}$$

Démonstration : L'implication directe est évidente : si

$$(d) = (a) + (b)$$

on a évidemment $(a) \subset (d)$ et $(b) \subset (d)$ ce qui montre que d est un diviseur commun à a et b. De plus, comme $d \in (d) = (a) + (b)$, l'existence de u et v est assurée. Réciproquement, si d est un diviseur commun à a et b, on a $(a) \subset (d)$ et $(b) \subset (d)$, donc, puisque (d) est stable pour l'addition, $(a) + (b) \subset (d)$. Si de plus, il existe $u, v \in \mathbb{A}$ avec $d = au + bv$, on a $d \in (a) + (b)$, donc $(d) \subset (a) + (b)$. La double inclusion montre bien que d est un PGCD de a et b. ∎

Conséquence : $d = \text{PGCD}\,(a,b)$ peut donc bien être interprété comme un "plus grand" diviseur commun : c'est un diviseur commun, et si δ en est un autre, δ divise tous les éléments de la forme $ax + by$, donc aussi $d = au + bv$. d **est donc un plus grand élément (pour la relation de divisibilité) parmi les diviseurs communs de a et b.** Cette affirmation n'est pas tout à fait correcte, puisque la relation de divisibilité n'est pas une relation d'ordre... Par contre, sur $\mathbb{Z}$, si l'on décide de choisir d positif, d est bien le plus grand diviseur commun de a et b au sens de la relation d'ordre de divisibilité sur $\mathbb{N}$ (donc aussi au sens de la relation d'ordre naturel, mais c'est moins précis). De même sur $\mathbb{K}[X]$, si $(A, B) \neq (0,0)$ et $D = \text{PGCD}\,(A,B)$, D est un multiple de tous les diviseurs communs à A et B, c'est donc aussi un diviseur commun de degré maximal, mais

ici encore, on perd de l'information si on ne garde que cet aspect de la définition de D.

De même, si on travaille avec une famille finie $(a_i)_{1\leqslant i\leqslant p}$ d'éléments de $\mathbb{A}$, on définira un PGCD par

$$d = \mathrm{PGCD}\,(a_i)_{1\leqslant i\leqslant p} \Leftrightarrow$$
$$(a_1, a_2, \ldots, a_p) = \{x \in \mathbb{A} \mid \exists x_1, \ldots x_p \in \mathbb{A} \quad x = a_1x_1 + \cdots + a_px_p\} = (d)$$

l'existence de d étant assuré par le caractère principal de $\mathbb{A}$. Ici encore, on peut caractériser d (à la relation "être associé à" près) par le fait que d est un diviseur commun à tous les $(a_i)_{1\leqslant i\leqslant p}$ pouvant s'écrire

$$d = a_1u_1 + \cdots + a_pu_p$$

Comme on a évidemment

$$(a_1, a_2, \ldots, a_p) = (a_1, a_2, \ldots, a_{p-1}) + (a_p)$$

il est aisé d'en déduire l'"**associativité**" de l'opération PGCD dans $\mathbb{A}$ (attention : ce n'est pas tout à fait une loi de composition interne, les égalités sont à prendre pour des équivalences pour l'association, remarque valable également pour l'énoncé suivant).

PROPOSITION 5-1.22 **Si** $a_1, a_2, \ldots, a_p$ **et** $b \in \mathbb{A}$

$$\mathrm{PGCD}\,(ba_1, ba_2, \ldots, ba_p) = b\,\mathrm{PGCD}\,(a_1, a_2, \ldots, a_p)$$

Démonstration : Il suffit de remarquer, avec des notations évidentes que

$$(ba_1, ba_2, \ldots, ba_p) = b\,(a_1, a_2, \ldots, a_p)$$

(égalité entre idéaux de $\mathbb{A}$), et par conséquent, si d est un générateur de l'idéal $(a_1, a_2, \ldots, a_p)$, bd engendre $(ba_1, ba_2, \ldots, ba_p)$. ■

5-1.2.2 Eléments premiers entre eux

DÉFINITION 5-1.23 *Des éléments* $(a_i)_{1\leqslant i\leqslant p}$ *de* $\mathbb{A}$ *sont dits* ***premiers entre eux*** *(dans leur ensemble) si et seulement si*

$$1_{\mathbb{A}} = \mathrm{PGCD}\,(a_1, a_2, \ldots, a_p)$$

Cela signifie que les seuls diviseurs communs à tous les a_i sont les unités de $\mathbb{A}$. Il ne faut donc pas confondre "premiers entre eux dans leur ensemble" et "premiers entre eux deux à deux" : il est clair que si $\mathrm{PGCD}\,(a_1, a_2) = 1_{\mathbb{A}}$, on a *a fortiori* $1_{\mathbb{A}} = \mathrm{PGCD}\,(a_1, a_2, \ldots, a_p)$. L'exemple (dans $\mathbb{Z}$) $a_1 = 6$, $a_2 = 10$ et $a_3 = 15$ montre que des entiers peuvent être premiers entre eux alors qu'aucun couple d'entre eux ne possède cette propriété.

Si p est un élément **irréductible** (*cf.* définition **5-1.**8) et si $a \in \mathbb{A}$, un PGCD de a et p est diviseur de p. A une association près, c'est donc $1_{\mathbb{A}}$ ou p. On a donc

PROPOSITION 5-1.24 Un élément irréductible de $\mathbb{A}$ est premier avec tout élément qu'il ne divise pas.

Les $(a_i)_{1\leqslant i\leqslant p} \in \mathbb{A}$ sont premiers entre eux si l'idéal qu'ils engendrent est égal à $(1_{\mathbb{A}}) = \mathbb{A}$. Comme un idéal est égal à l'anneau si et seulement s'il contient l'élément unité, on obtient immédiatement **le théorème de Bézout** :

THÉORÈME 5-1.25 Les éléments $(a_i)_{1\leqslant i\leqslant p}$ de $\mathbb{A}$ sont premiers entre eux si et seulement si

$$\exists\, u_1, \ldots, u_p \in \mathbb{A} \qquad a_1u_1 + \cdots + a_pu_p = 1_{\mathbb{A}}$$

Une utilisation du théorème **5-1**.22 donne immédiatement le résultat suivant, très souvent utilisé pour "simplifier par un PGCD" :

PROPOSITION 5-1.26 Si $(a_i)_{1\leqslant i\leqslant p} \in \mathbb{A}$, on a

$$d = \mathrm{PGCD}\,(a_1, a_2, \ldots, a_p) \quad \Longleftrightarrow$$

$$\exists\, a'_1, \ldots, a'_p \in \mathbb{A} \text{ **premiers entre eux avec** } \quad \begin{cases} a_1 = da'_1 \\ \vdots \\ a_p = da'_p \end{cases}$$

Le **théorème de Gauss** est conséquence du théorème de Bézout :

THÉORÈME 5-1.27 Si a, b et $c \in \mathbb{A}$, avec a et b premiers entre eux, et si a divise le produit bc, alors a divise c.

Démonstration : On écrit une identité de Bézout entre a et b :

$$\exists\, u, v \in \mathbb{A} \quad au + bv = 1_{\mathbb{A}}$$

En multipliant par c on obtient $acu + (bc)v = c$, ce qui montre que a divise c, puisqu'il divise à la fois acu et bcv. ■

COROLLAIRE 5-1.28 Si a et $b \in \mathbb{A}$ sont deux éléments premiers entre eux qui divisent un même élément $c \in \mathbb{A}$, le produit ab divise c.

Démonstration : c peut s'écrire ad, et est divisible par b. Le théorème de Gauss montre que b divise d, donc ab divise $ad = c$. ■

EXERCICE 5-1.29 Si $\mathrm{PGCD}\,(a, b) = 1_{\mathbb{A}}$, et si on connaît un couple (u_0, v_0) avec

$$au_0 + bv_0 = 1_{\mathbb{A}}$$

montrer que les couples (u, v) vérifiant une telle identité de Bézout sont exactement les couples s'écrivant

$$u = u_0 + xb \quad \text{et} \quad v = v_0 - xa \quad \text{avec} \quad x \in \mathbb{A} \text{ arbitraire}$$

Il résulte immédiatement du théorème de Gauss et de la proposition **5-1**.24 l'énoncé :

THÉORÈME 5-1.30 Dans un anneau principal[2], si un élément irréductible divise un produit de facteurs, il divise nécessairement au moins un des facteurs.

L'hypothèse d'irréductibilité est évidemment indispensable : dans $\mathbb{Z}$, 6 divise 4×3. Une autre conséquence du théorème de Bézout est :

THÉORÈME 5-1.31 Si a est premier avec chacun des $(a_i)_{1 \leqslant i \leqslant p} \in \mathbb{A}$, il est premier avec leur produit.

Démonstration : On raisonne par récurrence sur p. On se rend alors compte que la démonstration pour $p = 2$ suffit. Si a est premier avec a_1 et a_2, en écrivant des identités de Bézout

$$au_1 + a_1 v_1 = au_2 + a_2 v_2 = 1_{\mathbb{A}}$$

et en multipliant ces égalités, on obtient

$$a\left(au_1 u_2 + u_1 a_2 v_2 + u_2 a_1 v_1\right) + \left(a_1 a_2\right) v_1 v_2 = 1_{\mathbb{A}}$$

identité qui prouve que $\mathrm{PGCD}\left(a, a_1 a_2\right) = 1_{\mathbb{A}}$. ∎

On en déduit en particulier que, si a et b sont premiers entre eux, il en est de même de toute puissance de a et de toute puissance de b.

5-1.2.3 Plus petit multiple commun

Si $(a_i)_{1 \leqslant i \leqslant p} \in \mathbb{A}$, l'idéal intersection $\bigcap_{i=1}^{p} (a_i)$ est l'ensemble des multiples communs à tous les a_i. Un quelconque de ses générateurs sera appelé un PPCM des $(a_i)_{1 \leqslant i \leqslant p}$:

$$m = \mathrm{PPCM}\left(a_1, \ldots, a_p\right) \Leftrightarrow (m) = (a_1) \cap \cdots \cap (a_p)$$

Ici encore, cette "égalité" est à prendre pour ce qu'elle signifie. m est un multiple commun aux $(a_i)_{1 \leqslant i \leqslant p} \in \mathbb{A}$. C'est un "plus petit" en ce sens qu'il divise tous les multiples communs aux $(a_i)_{1 \leqslant i \leqslant p}$. Il y a "associativité" de l'opération. Ici encore il est clair que

$$\mathrm{PPCM}\left(ba_1, \ldots, ba_p\right) = b\,\mathrm{PPCM}\left(a_1, \ldots, a_p\right)$$

Il y a une relation simple entre PGCD et PPCM de deux éléments :

THÉORÈME 5-1.32 Si a et b sont premiers entre eux , leur PPCM est "égal" à leur produit. Plus généralement , pour a et b quelconques :

$$\mathrm{PGCD}(a, b) \times \mathrm{PPCM}(a, b) = a \times b$$

[2]Dans un anneau commutatif intègre, un élément p est dit **premier** s'il possède la propriété

$$p \text{ divise } xy \Rightarrow p \text{ divise } x \text{ ou } y$$

Il est facile de voir qu'un élément premier est toujours irréductible. Dans un anneau principal, il y a équivalence entre premier et irréductible.

Démonstration : Si m est un PPCM de a et b, c'est un diviseur de leur multiple commun ab. Si a et b sont premiers entre eux, le corollaire **5-1**.28 montre que ab divise m et, finalement, m et ab sont associés.

Dans le cas général, si $d = \text{PGCD}(a,b)$, on écrit $a = da'$ et $b = db'$ avec $\text{PGCD}(a',b') = 1_{\mathbf{A}}$. On alors

$$\text{PPCM}(a,b) = d \times \text{PPCM}(a',b') = da'b'$$

ce qui donne immédiatement le résultat. ■

Plus généralement, si les $(a_i)_{1 \leqslant i \leqslant p} \in \mathbb{A}$ sont **premiers entre eux deux à deux**, leur PPCM est égal à leur produit. Mais dès que $p \geqslant 3$, il n'existe plus dans le cas général de relation simple entre PPCM, PGCD et produit.

5-1.2.4 Algorithme d'Euclide

Dans un anneau comme $\mathbb{Z}$ ou $\mathbb{K}[X]$, où existe une division euclidienne, l'algorithme d'Euclide donne un moyen de calculer un PGCD de deux éléments. Rappelons ce résultat dans le cas de $\mathbb{K}[X]$:

Si A et $B \in \mathbb{K}[X]$ avec $B \neq 0$, l'identité de division euclidienne de A par B s'écrit

$$A = BQ + R \quad \text{avec} \quad R = 0 \text{ ou } d^\circ R < d^\circ B$$

Si $R = 0$ alors B divise A et donc $\text{PGCD}(A,B) = B$, sinon il est facile de voir que les diviseurs communs à A et B sont exactement les diviseurs communs à B et R, donc que

$$\text{PGCD}(A,B) = \text{PGCD}(B,R)$$

En réitérant l'opération, la suite des degrés des restes est strictement décroissante, et on atteint nécessairement une étape pour laquelle le reste est nul. **Le PGCD de A et B est alors le dernier reste non nul dans la suite des divisions euclidiennes ainsi effectuée** (algorithme d'Euclide).

Outre un moyen de calculer effectivement un PGCD [3], l'algorithme d'Euclide est aussi un moyen d'obtenir des polynômes U et V vérifiant

$$AU + BV = \text{PGCD}(A,B)$$

Il suffit de remarquer que chacun des restes de l'algorithme d'Euclide est dans l'idéal $(A) + (B)$, et qu'on peut, pour chacun de ces restes, obtenir de proche en proche une écriture sous la forme $A \times ? + B \times ??$. Si on démarre l'algorithme, on a par exemple

$$A = BQ_1 + R_1, \text{ puis } B = R_1Q_2 + R_2, \quad R_1 = R_2Q_3 + R_3 \ldots$$

[3] Avec une difficulté liée cependant à la méthode : en travaillant avec des polynômes "raisonnables" de $\mathbb{Q}[X]$, il se peut que les tailles des coefficients des restes intermédiaires de l'algorithme (numérateur et dénominateur de chacun des coefficients) croissent de manière très rapide... Les calculs doivent être menés en arithmétique exacte : dans $\mathbb{Q}[X]$, le PGCD de $X+1$ et de X^2-1 est $X+1$. Celui de $X+1$ et de $X^2-(1+10^{-50})$ vaut 1 !

ce qui donnera

$$\begin{cases} R_1 = A + B(-Q_1) \\ R_2 = B - R_1Q_2 = A(-Q_2) + B(1 + Q_2Q_1) \\ R_3 = R_1 - R_2Q_3 = A(1 + Q_3Q_2) + B(-Q_1 - Q_3 - Q_3Q_2Q_1) \quad \text{etc...} \end{cases}$$

Enfin, un résultat théorique important est l'**invariance du PGCD (et donc du PPCM) par extension du corps de base :**

Si A et B sont deux polynômes de $\mathbb{K}[X]$, et si $\mathbb{L}$ est un corps commutatif contenant $\mathbb{K}$, le PGCD de A et B sera le même (à association près...) s'il est évalué dans $\mathbb{K}[X]$ ou dans $\mathbb{L}[X]$, tout simplement parce que l'algorithme d'Euclide se déroulera entièrement dans $\mathbb{K}[X]$.

Par contre, la notion d'irréductibilité dans un anneau de polynômes dépend du corps de base : dans $\mathbb{R}[X]$, le polynôme $X^2 + 1$ est irréductible, il ne l'est pas dans $\mathbb{C}[X]$. Cela doit faire réfléchir au résultat de la section suivante, où l'on montre que dans $\mathbb{K}[X]$ (et, pour les curieux, dans tout anneau principal) PGCD et PPCM peuvent s'exprimer à l'aide de diviseurs irréductibles.

5-1.2.5 Décomposition en facteurs irréductibles

Rappelons le résultat du programme de première année :

THÉORÈME 5-1.33 Tout polynôme de degré au moins 1 de $\mathbb{K}[X]$ admet une décomposition comme produit d'un scalaire non nul et de polynômes irréductibles. Cette décomposition est unique à l'ordre des facteurs près si on impose à ces diviseurs irréductibles d'être normalisés.

Revoyons la démarche de la démonstration : si $P = P_0$ est irréductible, il suffit de le normaliser. Sinon on peut lui trouver un diviseur irréductible I_1 de la manière suivante : P possède un diviseur strict Q_1. Si Q_1 est irréductible, on prend $I_1 = Q_1$. Sinon on choisit un diviseur strict de Q_1, soit Q_2 et on continue la démarche ... Comme les degrés des Q_i forment une suite d'entiers strictement décroissante, on arrive nécessairement au bout d'une nombre fini d'étapes à un diviseur irréductible. On recommence avec le polynôme $P_1 = P_0/I_1$ etc... Ici encore le processus s'arrête car les degrés des P_i vont en décroissant. L'unicité se démontre à l'aide du théorème de Gauss.

En changeant de notations et en regroupant les facteurs irréductibles apparaissant plusieurs fois dans la décomposition de P, on obtiendra une décomposition unique (à l'ordre près) sous la forme

$$P = \lambda P_1^{\alpha_1} \cdots P_k^{\alpha_k}$$

avec $\lambda \in \mathbb{K}^*$, les P_i polynômes irréductibles normalisés distincts deux à deux et $\alpha_i \in \mathbb{N}^*$.

EXERCICE 5-1.34 Montrer que, si $\mathbb{K}$ est un corps fini, il existe dans $\mathbb{K}[X]$ des polynômes irréductibles dont le degré est plus grand qu'un entier choisi arbitrairement.

COROLLAIRE 5-1.35 Si A et $B \in \mathbb{K}[X]$ sont de degré au moins 1, B est un diviseur de A ssi tous les diviseurs irréductibles apparaissant dans la

décomposition de B apparaissent aussi dans la décomposition de A, avec un exposant au moins égal.

COROLLAIRE 5-1.36 Si A et $B \in \mathbb{K}[X]$ sont de degré au moins 1, si $P_1, P_2, \ldots, P_k$ sont les polynômes irréductibles normalisés apparaissant dans au moins une des décompositions de A et B :

$$A = \lambda P_1^{\alpha_1} \cdots P_k^{\alpha_k} \quad \textbf{et} \quad B = \mu P_1^{\beta_1} \cdots P_k^{\beta_k}$$

(avec cette fois $\alpha_i \in \mathbb{N}$, $\beta_i \in \mathbb{N}$ et $\alpha_i + \beta_i > 0$), on a

$$\begin{cases} \mathrm{PGCD}(A, B) = P_1^{\min(\alpha_1, \beta_1)} \cdots P_k^{\min(\alpha_k, \beta_k)} \\ \mathrm{PPCM}(A, B) = P_1^{\max(\alpha_1, \beta_1)} \cdots P_k^{\max(\alpha_k, \beta_k)} \end{cases}$$

(on retrouve évidemment le résultat du théorème 5-1.32)

Les résultats énoncés ci-dessus se retrouvent dans $\mathbb{Z}$, pour peu que l'on remplace

"polynôme de degré > 0"	par	"entier $\in \mathbb{Z} - \{0, \pm 1\}$"
"scalaire non nul"	par	"± 1"
"polynôme irréductible normalisé"	par	"entier naturel premier"

les démonstrations sont identiques. Qu'en est-il dans un anneau principal quelconque ?

EXERCICE 5-1.37 Montrer que les résultats précédents sont vrais dans un anneau principal (existence d'une décomposition en facteurs irréductibles, unicité à l'ordre des facteurs et aux unités près, expression du PGCD et du PPCM). Pour paraphraser la démonstration du théorème **5-1.33**, on démontrera d'abord le

LEMME 5-1.38 Dans un anneau principal, toute suite croissante d'idéaux (pour l'inclusion) est constante à partir d'un certain rang.

5-2 Calculs polynomiaux dans une algèbre

Si $(\mathbb{A}, +, \times, .)$ est une $\mathbb{K}$-algèbre (*cf.* **2-2.3**), on peut, pour $a \in \mathbb{A}$, définir une combinaison linéaire de la famille $(a^i)_{i \in \mathbb{N}}$. (On rappelle que, par convention, on pose $a^0 = 1_{\mathbb{A}}$). Cela revient en fait, dans un polynôme $P \in \mathbb{K}[X]$, à substituer à l'indéterminée X la "valeur" $a \in \mathbb{A}$.

5-2.1 Morphisme $P \mapsto P(a)$. Algèbre $\mathbb{K}[a]$

DÉFINITION 5-2.1 *Si $\mathbb{A}$ est une $\mathbb{K}$-algèbre et $P = \sum_{k=0}^{m} \alpha_k X^k \in \mathbb{K}[X]$, pour $a \in \mathbb{A}$, on définit $P(a) \in \mathbb{A}$ par*

$$P(a) = \sum_{k=0}^{m} \alpha_k a^k = \alpha_0 1_{\mathbb{A}} + \alpha_1 a + \cdots + \alpha_m a^m$$

On notera que l'on remplace X^0 (polynôme habituellement identifié à $1_{\mathbb{K}}$) par $1_{\mathbb{A}}$ et pour $k \geqslant 1$, X^k par a^k dans le polynôme P. Les règles de calcul dans une $\mathbb{K}$-algèbre (distributivité de la multiplication par rapport à l'addition, exportativité des scalaires etc...) montrent facilement que

THÉORÈME 5-2.2 Si a est un élément fixé d'une $\mathbb{K}$-algèbre $(\mathbb{A}, +, \times, .)$, l'application

$$\mathbb{K}[X] \to \mathbb{A} \quad P \mapsto P(a)$$

est un morphisme de $\mathbb{K}$-algèbres (cf. définition 2-2.7).

Démonstration : Les vérifications des égalités

$$(P + \lambda Q)(a) = P(a) + \lambda Q(a) \text{ et } (PQ)(a) = P(a)Q(a)$$

sont immédiates. ■

Certains exemples de telles applications sont bien connus : dans le cas $\mathbb{A} = \mathbb{K}$, il s'agit d'une évaluation d'un polynôme sur un élément du corps de base, notion derrière celle de fonction polynôme. Le cas $\mathbb{A} = \mathbb{K}[X]$ correspond à la notion de **composition de polynômes** : si $A \in \mathbb{K}[X]$, on a simplement $P(A) = P \circ A$.

Nous étudierons plus particulièrement les cas $\mathbb{A} = \mathcal{L}_{\mathbb{K}}(\mathbb{E})$, algèbre des endomorphismes d'un $\mathbb{K}$-ev $\mathbb{E}$ et $\mathbb{A} = \mathcal{M}_n(\mathbb{K})$ algèbre des matrices carrées à coefficients dans $\mathbb{K}$. Ce second exemple est la copie du premier, lorsque $\mathbb{E}$ est de dimension finie n, pour peu que l'on fixe une base de $\mathbb{E}$. On parle alors de polynômes d'endomorphismes ou de matrices.

THÉORÈME 5-2.3 (et définition) L'image du morphisme

$$\mathbb{K}[X] \to \mathbb{A} \quad P \mapsto P(a)$$

est une sous-algèbre de $\mathbb{A}$, qu'on notera $\mathbb{K}[a]$. C'est clairement la plus petite sous-algèbre de $\mathbb{A}$ contenant a. On l'appelle donc sous-algèbre engendrée par a. Elle est commutative.

Démonstration : évident.

On a donc clairement

$$\mathbb{K}[a] = \{P(a),\ P \in \mathbb{K}[X]\} = \text{vect}\left(a^i\right)_{i \in \mathbb{N}}$$

5-2.2 Idéal annulateur

Dans l'algèbre $\mathcal{M}_3(\mathbb{R})$, considérons la matrice

$$A = \begin{pmatrix} -9 & -10 & -20 \\ 4 & 5 & 8 \\ 2 & 2 & 5 \end{pmatrix}$$

et supposons que l'on ait à calculer $P(A)$, avec $P = 27X^{15} + X^7 - 2X^4 + 3$. Ceci peut sembler fort calculatoire au début, mais si l'on calcule A^2, on trouve I_2. Il en résulte évidemment que les puissances impaires de A valent A, les puissances paires valant évidemment I_2. On a donc

$$P(A) = 28A + I_2$$

dont la valeur numérique s'obtient finalement presque sans calcul. Ce genre de simplification se généralise. Si on cherche à formaliser le calcul qu'on vient de faire, on a remarqué que le polynôme

$$Q = X^2 - 1$$

vérifie $Q(A) = 0$ (on traduit exactement que A est matrice d'une involution, c'est-à-dire d'une symétrie vectorielle). Remplacer dans $P(A)$ les puissances de A par A ou I_2 selon la parité de l'exposant revient exactement à calculer $R(A)$, où R est le reste de la division euclidienne de P par Q. On généralise avec les définitions et résultats suivants :

DÉFINITION 5-2.4 *Si a est un élément d'une $\mathbb{K}$-algèbre $(\mathbb{A}, +, \times, .)$, un polynôme $P \in \mathbb{K}[X]$ est dit **annulateur** pour a (on dit aussi a annule P) ssi $P(a) = 0_{\mathbb{A}}$.*

Lorsque $\mathbb{A} = \mathbb{K}$ (ou un sur-corps de $\mathbb{K}$), on peut dire aussi que a est une racine de P. **On réservera la terminologie "racine" à cette situation particulière,** en particulier pour éviter des erreurs concernant le nombre d'éléments d'une algèbre annulés par un polynôme : dans $\mathcal{M}_3(\mathbb{R})$, le polynôme $X^2 - 1$ annule une infinité de matrices. Les solutions de $A^2 - I_2$ vérifient

$$(A - I_2)(A + I_2) = 0_3$$

mais l'algèbre $\mathcal{M}_3(\mathbb{R})$ n'est pas intègre. Le polynôme $X^2 - 1$ n'a évidemment que deux racines réelles, $\mathbb{R}$ étant un corps.

THÉORÈME 5-2.5 (et définition) Si a appartient à une $\mathbb{K}$-algèbre $(\mathbb{A}, +, \times, .)$, le noyau du morphisme $P \mapsto P(a)$ est un idéal de $\mathbb{K}[X]$. C'est l'idéal formé de tous les polynômes annulateurs de a. On l'appelle idéal annulateur de a.

On a alors deux possibilités :

- Le morphisme $P \mapsto P(a)$ est injectif, son noyau étant réduit à (0). a ne possède alors pas de polynôme annulateur non trivial, et l'algèbre $\mathbb{K}[a]$ est isomorphe à $\mathbb{K}[X]$. En d'autres termes, si on effectue des calculs sur des polynômes en a, il n'y a pas moyen d'abaisser les degrés en faisant des simplifications comme nous avons pu le faire sur l'exemple introductif de cette section.
- L'idéal annulateur de a n'est pas réduit à zéro. C'est donc un idéal **principal** de $\mathbb{K}[X]$, engendré par un élément de $\mathbb{K}[X]$, polynôme annulateur qui divise tous les autres, donc en particulier de degré minimal.

DÉFINITION 5-2.6 *Si a est un élément d'une $\mathbb{K}$-algèbre $(\mathbb{A},+,\times,.)$ qui possède un idéal annulateur non trivial, on dit aussi que a possède un polynôme minimal, et plus précisément on appelle* ***polynôme minimal de*** *a tout générateur de cet idéal annulateur.*

Si π est un polynôme minimal de a, on a donc

$$\forall P \in \mathbb{K}[X] \quad P(a) = 0_{\mathbb{A}} \Leftrightarrow \pi \text{ divise } P$$

Remarquons qu'avec notre définition, il n'y a pas unicité du polynôme minimal, il s'agit plutôt d'une classe d'équivalence pour la relation "être associé à". Il y aurait bien sûr unicité si on imposait à π d'être normalisé. (On se permettra cependant l'abus de langage "soit π **le** polynôme minimal de a"...). Par exemple il est facile de voir que X^2-1 est un polynôme minimal pour la matrice $A \in \mathcal{M}_3(\mathbb{R})$ envisagée au début de cette section. Réciproquement, si $B \in \mathcal{M}_3(\mathbb{R})$ vérifie

$$B^2 = I_2$$

le polynôme X^2-1 est annulateur pour B. Comme il ne possède (à un scalaire multiplicatif non nul près) que les diviseurs stricts $X-1$ et $X+1$, on a trois possibilités :

- Le polynôme minimal π_B vaut $X-1$ et on a $B = I_2$.
- $\pi_B = X+1$ et donc $B = -I_2$.
- $\pi_B = X^2-1$, et B est matrice d'une symétrie vectorielle de $\mathbb{R}^3$ par rapport à un sev $\mathbb{E}_1$ parallèlement à un sev $\mathbb{E}_{-1}$ de dimensions strictement positives. Cela signifie que B est semblable à une des matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ ou } \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

L'ensemble des matrices de $\mathcal{M}_3(\mathbb{R})$ annulées par X^2-1 est donc réunion de quatre classes de similitudes.

THÉORÈME 5-2.7 Un élément a d'une $\mathbb{K}$-algèbre $(\mathbb{A},+,\times,.)$ possède un polynôme minimal π_a si et seulement si l'algèbre $\mathbb{K}[a]$ est de dimension finie. On a alors

$$\dim \mathbb{K}[a] = d^\circ \pi_a$$

et plus précisément, si d est ce degré, $\{1_{\mathbb{A}}, a, \ldots, a^{d-1}\}$ est une base de $\mathbb{K}[a]$.

Démonstration : Si a n'a pas de polynôme minimal, $\mathbb{K}[a]$ est isomorphe à $\mathbb{K}[X]$ et est donc de dimension infinie. Si à l'inverse a possède un polynôme minimal π_a de degré d, il est clair que la famille $\{1_{\mathbb{A}}, a, \ldots, a^{d-1}\}$ est libre dans $\mathbb{K}[a]$, car une relation de liaison non triviale

$$\sum_{k=0}^{d-1} \alpha_k a^k = 0_{\mathbb{A}}$$

donnerait un polynôme annulateur non nul $P = \sum_{k=0}^{d-1} \alpha_k X^k$ de degré strictement inférieur à celui de π_a, en contradiction avec la définition de π_a. De plus, la famille $\{1_{\mathbb{A}}, a, \dots, a^{d-1}\}$ engendre $\mathbb{K}[a]$, puisque, si $b \in \mathbb{K}[a]$, il existe un polynôme $P \in \mathbb{K}[X]$ avec $b = P(a)$. L'identité de division euclidienne de P par π_a s'écrit

$$P = \pi_a Q + R \quad \text{avec } R = 0 \text{ ou } d^\circ R \leqslant d - 1$$

ce qui donne bien

$$b = P(a) = \pi_a(a) Q(a) + R(a) = R(a) \in \text{vect}\{1_{\mathbb{A}}, a, \dots, a^{d-1}\} \quad ■$$

EXERCICE 5-2.8 Si $a \in (\mathbb{A}, +, \times, .)$ possède un polynôme minimal, montrer qu'il en est de même pour tout élément de $\mathbb{A}$ pouvant s'écrire $P(a)$, où $P \in \mathbb{K}[X]$.

COROLLAIRE 5-2.9 Dans une $\mathbb{K}$-algèbre de dimension finie, tout élément possède un polynôme minimal.

C'est en particulier le cas pour l'algèbre des matrices carrées $\mathcal{M}_n(\mathbb{K})$ ou l'algèbre des endomorphismes d'un espace vectoriel de dimension finie. Si $\dim \mathbb{A} = p$, pour tout $a \in \mathbb{A}$ on aura $\dim \mathbb{K}[a] \leqslant \dim \mathbb{A} = p$, donc le polynôme minimal de a est de degré[4] au maximum égal à p.

EXERCICE 5-2.10 Soit $Q \in \mathbb{K}[X]$. A quelle condition l'endomorphisme de $\mathbb{K}[X]$ défini par $P \mapsto QP$ possède-t-il un polynôme minimal ?

EXERCICE 5-2.11 Soit $\mathbb{L}$ un sur-corps de $\mathbb{K}$ et $A \in \mathcal{M}_n(\mathbb{K})$. Montrer qu'un polynôme minimal $\pi \in \mathbb{K}[X]$ de $A \in \mathcal{M}_n(\mathbb{K})$ est aussi polynôme minimal de A considérée comme matrice de $\mathcal{M}_n(\mathbb{L})$.

Ce qu'il faut retenir de ce qui précède est essentiellement ceci :
Dans la sous-algèbre $K[a]$ de $(\mathbb{A}, +, \times, .)$ engendrée par un élément a possédant un polynôme minimal π, les calculs peuvent être menés "modulo π", c'est-à-dire que, pour évaluer $P(a)$, on peut, pour abaisser le degré, effectuer d'abord la division euclidienne de P par π, et calculer $R(a)$ où R est le reste de cette division. Il est d'ailleurs à remarquer qu'on peut de même abaisser le degré dès qu'on connaît un polynôme annulateur (non nécessairement "minimal").

EXERCICE 5-2.12 Calculer $P\left(1 + j\sqrt[3]{2} + j^2\sqrt[3]{4}\right)$ où

$$j = e^{2i\pi/3} \text{ et } P = X^4 - 5X^3 + 3X^2 + 18X - 21$$

EXERCICE 5-2.13 Si $A \in \mathcal{M}_n(\mathbb{R})$ vérifie $A^2 - 5A = 6I_n$, expliciter A^{1998} comme combinaison linéaire de A et I_n.

Si a possède un polynôme minimal, il est facile de caractériser les éléments inversibles de l'algèbre $\mathbb{K}[a]$:

[4]Dans le cas de l'algèbre $\mathcal{M}_n(\mathbb{K})$, pour laquelle $p = n^2$, on verra mieux plus loin (théorème de Cayley-Hamilton) : le polynôme minimal d'une matrice de $\mathcal{M}_n(\mathbb{K})$ est de degré au maximum égal à n.

THÉORÈME 5-2.14 Si un élément a d'une $\mathbb{K}$-algèbre $(\mathbb{A},+,\times,.)$ possède un polynôme minimal π_a, un élément $P(a)$ est une unité de $\mathbb{A}$ si et seulement si le polynôme P est premier avec π_a. L'inverse de $P(a)$ est alors lui-même dans $\mathbb{K}[a]$, donc peut s'écrire comme un polynôme en a.

Démonstration : Si P est premier avec π_a, on peut écrire une identité de Bézout entre P et π_a :

$$\exists U, V \in \mathbb{K}[X] \quad UP + V\pi_a = 1$$

ce qui donne

$$U(a)P(a) = 1_{\mathbb{A}}$$

et prouve l'inversibilité de $P(a)$ avec $(P(a))^{-1} = U(a) \in \mathbb{K}[a]$ [5]. Si P n'est pas premier avec π_a, en notant $D = \mathrm{PGCD}(P, \pi_a)$, on peut écrire

$$P = DP_1 \text{ et } \pi_a = D\pi^1$$

et comme $d^\circ D > 0$, on a $\pi^1(a) \neq 0_{\mathbb{A}}$ (par définition du polynôme minimal). On a cependant

$$\pi^1(a)P(a) = \pi^1(a)D(a)P_1(a) = \pi_a(a)P_1(a) = 0_{\mathbb{A}}$$

ce qui montre que $P(a)$ est diviseur de zéro dans $\mathbb{A}$ et ne peut pas, par conséquent, être inversible. ■

En particulier a est inversible ssi X ne divise pas π_a, c'est-à-dire ssi $\pi_a(0) \neq 0$. Le calcul de l'inverse de a est alors immédiat : par exemple, si $A \in \mathcal{M}_n(\mathbb{K})$ vérifie $A^3 - 2A^2 + 5I_n = 0_n$, on a immédiatement

$$5I_n = -A(A^2 - 2A) \Rightarrow A^{-1} = \frac{1}{5}(2A - A^2)$$

EXERCICE 5-2.15 Si a n'a pas de polynôme minimal, quels sont les éléments de l'algèbre $\mathbb{K}[a]$ inversibles dans $\mathbb{K}[a]$?

5-2.3 Polynômes d'endomorphismes et matrices

Si $\mathbb{E}$ est un $\mathbb{K}$-ev de dimension finie n, et si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$, lorsqu'on représente u par sa matrice A dans une base $\mathcal{B}$, on a évidemment, pour $P \in \mathbb{K}[X]$

$$M(P(u), \mathcal{B}) = P(A)$$

Cela montre aussi un résultat qui est algébriquement évident (toute similitude associée à une matrice inversible est un automorphisme de l'algèbre $\mathcal{M}_n(\mathbb{K})$) :

Pour $A \in \mathcal{M}_n(\mathbb{K})$, $M \in \mathcal{GL}_n(\mathbb{K})$ et $P \in \mathbb{K}[X]$

$$P(M^{-1}AM) = M^{-1}P(A)M$$

Les calculs polynomiaux "commutent" d'une certaine manière avec les similitudes de $\mathcal{M}_n(\mathbb{K})$ [6].

[5] Voir également l'exercice **4-1**.30

[6] Plus généralement, tout automorphisme d'une algèbre $\mathbb{A}$ commute avec les calculs polynomiaux

EXERCICE 5-2.16 Montrer qu'une matrice et sa transposée ont même polynôme minimal.

5-2.4 Cas d'une algèbre intègre

Nous envisageons ici le cas particulier où la $\mathbb{K}$-algèbre $\mathbb{A}$ est **intègre**[7]

THÉORÈME 5-2.17 Si $(\mathbb{A}, +, \times, .)$ est intègre et si $a \in \mathbb{A}$ possède un polynôme minimal, celui-ci est irréductible.

Démonstration : Si π_a n'est pas irréductible, il admet une factorisation en diviseurs non triviaux

$$\pi_a = \pi^1 \pi^2$$

et on a alors, pour des raisons de degré

$$\pi^1(a) \neq 0_{\mathbb{A}},\ \pi^2(a) \neq 0_{\mathbb{A}} \text{ et } \pi^1(a)\pi^2(a) = 0_{\mathbb{A}}$$

ce qui nie l'intégrité de $\mathbb{A}$. ■

COROLLAIRE 5-2.18 Si $(\mathbb{A}, +, \times, .)$ est intègre, un élément $a \in \mathbb{A}$ possède un polynôme minimal si et seulement si $(\mathbb{K}[a], +, \times)$ est un corps.

Démonstration : Si a n'a pas de polynôme minimal $(\mathbb{K}[a], +, \times)$ est isomorphe à $(\mathbb{K}[X], +, \times)$ et n'est pas un corps. Si a possède un polynôme minimal π_a, celui-ci est irréductible. Si $P(a)$ est un élément non nul de $\mathbb{K}[a]$, π_a ne divise pas P, donc est premier avec P, et donc $P(a)$ est inversible dans $\mathbb{K}[a]$ (théorème **5-2**.14). ■

EXERCICE 5-2.19 Si $\mathbb{A}$ n'est pas intègre, à quelle condition nécessaire et suffisante $\mathbb{K}[a]$ est-il un corps ?

5-3 Sous-espace stable par un endomorphisme

Dans le reste de ce chapitre, nous utiliserons des polynômes d'endomorphismes (ou de matrices) pour simplifier l'étude de ces endomorphismes, essentiellement pour "réduire les dimensions" des espaces sur lesquels ils opèrent.

[7] Ce n'est évidemment pas le cas dans ce qui nous intéressera dans la suite de ce chapitre, à savoir les algèbres de matrices carrées ou d'endomorphismes d'un espace vectoriel. Par contre, un cas théoriquement très important où les résultats de cette section s'appliquent est celui où $\mathbb{A} = \mathbb{L}$ est un sur-corps de $\mathbb{K}$. Un élément a de $\mathbb{L}$ possédant un polynôme minimal dans $\mathbb{K}[X]$ est alors dit **algébrique** sur $\mathbb{K}$. Il est dit **transcendant** sur $\mathbb{K}$ dans le cas contraire.

5-3.1 Endomorphisme induit sur un sous-espace vectoriel stable

Soit $\mathbb{E}$ un $\mathbb{K}$-ev et $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$. Un sous-espace vectoriel $\mathbb{F}$ de $\mathbb{E}$ est dit **stable par** u (en abrégé u-stable) si et seulement si

$$u(\mathbb{F}) \subset \mathbb{F}$$

Il est alors clair qu'on peut considérer (en commettant un petit abus de langage, correspondant à une restriction de l'espace d'arrivée) que $u|_{\mathbb{F}}$ est une application linéaire de $\mathbb{F}$ dans $\mathbb{F}$, qu'on appelle **endomorphisme de $\mathbb{F}$ induit**[8] **par** u.

Si l'espace $\mathbb{E}$ est de **dimension finie** n, on sait qu'on peut représenter u par sa matrice dans une base $\mathcal{B}$ de $\mathbb{E}$. Si cette base est choisie n'importe comment, il est peut être difficile de voir qu'un sous-espace vectoriel $\mathbb{F}$ de dimension p est u-stable. Par contre, si **la base $\mathcal{B}$ est adaptée à $\mathbb{F}$** (c'est-à-dire que les p premiers vecteurs de $\mathcal{B}$ forment une base de $\mathbb{F}$, les $n-p$ derniers formant une base d'un supplémentaire $\mathbb{S}$ de $\mathbb{F}$), on a une caractérisation de la stabilité de $\mathbb{F}$ par u :

$\mathbb{F}$ sera u-stable si et seulement si les images des vecteurs d'une base de $\mathbb{F}$ sont dans $\mathbb{F}$, ce qui donne :

PROPOSITION 5-3.1 Si $\mathbb{F}$ est un sous-espace de dimension p d'un espace $\mathbb{E}$ de dimension n, et $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$. $\mathbb{F}$ est u-stable si et seulement s'il existe une base de $\mathbb{E}$ adaptée à $\mathbb{F}$ dans laquelle la matrice de u se décompose par blocs

$$\mathcal{M}(u, \mathcal{B} = \mathcal{B}_{\mathbb{F}} \cup \mathcal{B}_{\mathbb{S}}) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

où l'on a bien sûr

$$A = \mathcal{M}(u|_{\mathbb{F}}, \mathcal{B}_{\mathbb{F}}) \in \mathcal{M}_p(\mathbb{K})$$

et donc

$$\det u = \det A \, \det C = \det u|_{\mathbb{F}} \, \det C$$

La matrice aura alors une forme analogue dans toute base de $\mathbb{E}$ adaptée à $\mathbb{F}$.

Si de plus $\mathbb{S}$ est lui-même u-stable, cela se traduira sur la matrice précédente par $B = 0$ (dans ce cas, on aura aussi $C = \mathcal{M}(u|_{\mathbb{S}}, \mathcal{B}_{\mathbb{S}}) \in \mathcal{M}_{n-p}(\mathbb{K})$), et on pourra interpréter u comme recollement des endomorphismes $u|_{\mathbb{F}} \in \mathcal{L}_{\mathbb{K}}(\mathbb{F})$ et $u|_{\mathbb{S}} \in \mathcal{L}_{\mathbb{K}}(\mathbb{S})$ [9]

---

[8] Cette notion n'a évidemment de sens que si $\mathbb{F}$ est u-stable. Dans le cas contraire, on a simplement $u|_{\mathbb{F}} \in \mathcal{L}_{\mathbb{K}}(\mathbb{F}, \mathbb{E})$.

[9] Plus généralement, si $\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{F}_i$, les sous-espaces $\mathbb{F}_i$ sont tous u-stables pour $i = 1, \ldots, p$ si et seulement si, dans toute base $\mathcal{B} = \bigcup_{i=1}^{p} \mathcal{B}_i$ adaptée à la décomposition de $\mathbb{E}$, la matrice de u

Cette situation est particulière : un sous-espace stable par un endomorphisme ne possède en général pas de supplémentaire stable. Par exemple, l'endomorphisme de $\mathbb{R}^2$ exprimé dans la base canonique par la matrice

$$M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

possède une unique droite vectorielle stable (laquelle ?) qui ne possède donc évidemment pas de supplémentaire stable.

Le théorème suivant est souvent utilisé pour déterminer des sous-espaces stables par un endomorphisme :

THÉORÈME 5-3.2 Si deux endomorphismes u et v d'un espace vectoriel $\mathbb{E}$ commutent, le noyau et l'image de v sont deux sous-espaces stables par u.

Démonstration : Soit $x \in \ker v$. On a

$$v(u(x)) = u(v(x)) = u(0_{\mathbb{E}}) = 0_{\mathbb{E}}$$

donc $u(x) \in \ker v$, qui est donc u-stable. De même, si $x \in \operatorname{Im} v$, il existe un vecteur $y \in \mathbb{E}$ avec $x = v(y)$. On a alors

$$u(x) = u(v(y)) = v(u(y)) \in \operatorname{Im} v$$

L'image de v est donc aussi u-stable. ■

Si $u \circ v = v \circ u$, on vérifie immédiatement par récurrence sur p que

$$\forall p \in \mathbb{N} \qquad v^p \circ u = u \circ v^p$$

et, par linéarité

$$\forall P \in \mathbb{K}[X] \qquad P(v) \circ u = u \circ P(v)$$

et en conséquence

$u \circ v = v \circ u \Rightarrow$ le noyau (ou l'image) de tout polynôme en v est u-stable

En particulier, comme u commute avec tout polynôme en u

$$\forall P \in \mathbb{K}[X] \qquad \ker P(u) \text{ et } \operatorname{Im} P(u) \text{ sont } u\text{-stables.}$$

---

est diagonale par blocs :

$$\mathcal{M}(u, \mathcal{B}) = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_p \end{pmatrix}$$

avec $A_i = \mathcal{M}(u|_{\mathbb{F}_i}, \mathcal{B}_i)$. On a alors

$$\det u = \prod_{i=1}^{p} \det u|_{\mathbb{F}_i}$$

5-3.2 Théorème de décomposition des noyaux

Ce théorème permet souvent de décomposer le noyau d'un polynôme en u en somme directes de noyaux d'autres polynômes en u : on décompose ainsi un sous-espace u-stable en somme de sous-espaces stables (en général de dimensions plus petites).

THÉORÈME 5-3.3 Si P et Q sont deux polynômes de $\mathbb{K}[X]$ premiers entre eux et $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$, on a

$$\ker PQ(u) = \ker P(u) \circ Q(u) = \ker P(u) \oplus \ker Q(u)$$

Démonstration : Comme $PQ(u) = P(u) \circ Q(u) = Q(u) \circ P(u)$, il est clair que $\ker P(u)$ et $\ker Q(u)$ sont deux sous-espaces vectoriels de $\ker PQ(u)$. Il s'agit de montrer qu'ils sont indépendants, et que leur somme (directe) est égale à $\ker PQ(u)$. Comme $\mathrm{PGCD}(P, Q) = 1$ on utilise une identité de Bézout

$$\exists U, V \in \mathbb{K}[X] \qquad PU + QV = 1$$

On en déduit

$$U(u)\,P(u) + V(u)\,Q(u) = id_{\mathbb{E}}$$

Si $x \in \ker P(u) \cap \ker Q(u)$ on a alors

$$x = U(u)\,P(u)(x) + V(u)\,Q(u)(x) = 0_{\mathbb{E}}$$

$\ker P(u)$ et $\ker Q(u)$ sont donc indépendants. De plus, si $x \in \ker PQ(u)$, on a

$$P(u)(x) \in \ker Q(u) \text{ et } Q(u)(x) \in \ker P(u)$$

et *a fortiori*

$$x_1 = U(u)\,P(u)(x) \in \ker Q(u) \text{ et } x_2 = V(u)\,Q(u)(x) \in \ker P(u)$$

La décomposition

$$x = U(u)\,P(u)(x) + V(u)\,Q(u)(x) = x_1 + x_2$$

montre que $x \in \ker P(u) \oplus \ker Q(u)$, ce qui démontre le résultat. ■

COROLLAIRE 5-3.4 Si $(P_i)_{i=1..p}$ sont des polynômes premiers entre eux deux à deux, on a

$$\ker \prod_{i=1}^{p} P_i(u) = \bigoplus_{i=1}^{p} \ker P_i(u)$$

Démonstration : Il suffit de raisonner par récurrence sur p, en utilisant le résultat du théorème **5-1.31**. ■

Le théorème précédent (et son corollaire) permettent, en particulier, d'obtenir une décomposition de l'espace en somme directe de sous-espaces stables par un endomorphisme u dès que l'on connaît un polynôme annulateur factorisé en facteurs premiers entre eux deux à deux :

COROLLAIRE 5-3.5 Si $P = P_1 P_2 \cdots P_m$ est un polynôme annulateur d'un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ avec

$$i \neq j \Rightarrow \text{PGCD}(P_i, P_j) = 1$$

on a alors

$$\mathbb{E} = \ker P_1(u) \oplus \ker P_2(u) \oplus \cdots \oplus \ker P_m(u)$$

REMARQUE 5-3.6 En reprenant la démonstration du théorème qui précède, on s'aperçoit aisément que les projecteurs associés à cette décomposition de l'espace sont alors tous dans l'algèbre $\mathbb{K}[u]$, et peuvent donc s'écrire comme polynômes en u. Par exemple, si on reprend une identité de Bézout entre P_1 et le produit $P_2 \cdots P_m$ on obtient

$$P_1 U(u) + P_2 \cdots P_m V(u) = id_{\mathbb{E}}$$

et on vérifie que le projecteur p_1 sur $\ker P_1(u)$ parallèlement à la somme directe $\ker P_2(u) \oplus \cdots \oplus \ker P_m(u)$ n'est autre que

$$p_1 = P_2 \cdots P_m V(u) \in \mathbb{K}[u]$$

EXEMPLE 5-3.7 Décrire géométriquement un endomorphisme $u \in \mathcal{L}_{\mathbb{R}}(\mathbb{E})$ vérifiant

$$3u - u^2 = 2id_{\mathbb{E}}$$

Le polynôme annulateur dont on dispose ici est $X^2 - 3X + 2 = (X-1)(X-2)$, factorisé en deux polynômes du premier degré premiers entre eux. On a donc

$$\mathbf{E} = \ker(u - id_{\mathbf{E}}) \oplus \ker(u - 2id_{\mathbf{E}})$$

Si l'un de ces deux sous-espace est réduit à $\{0_{\mathbf{E}}\}$, on obtient le cas où u est une homothétie, avec plus précisément

$$u = id_{\mathbb{E}} \text{ ou } u = 2id_{\mathbb{E}}$$

Dans le cas général, il existe deux sous-espaces supplémentaires non triviaux $\mathbb{E}_1 = \ker(u - id_{\mathbb{E}})$ et $\mathbb{E}_2 = \ker(u - 2id_{\mathbb{E}})$ sur lesquels u opère respectivement comme des homothéties de rapport 1 et 2 : u peut donc être considéré comme recollement de deux endomorphismes simples des sous-espaces $\mathbb{E}_1$ et $\mathbb{E}_2$. Cela permet notamment de simplifier les calculs sur u : si $x \in \mathbb{E}$ est décomposé en $x = x_1 + x_2$ dans la somme directe $\mathbf{E}_1 \oplus \mathbb{E}_2$, on aura évidemment

$$\forall p \in \mathbb{Z} \quad u^p(x) = x_1 + 2^p x_2$$

De même, si l'espace est de dimension finie, dans une base de $\mathbb{E}$ adaptée à la décomposition $\mathbb{E}_1 \oplus \mathbb{E}_2$, la matrice de u sera particulièrement simple, puisqu'elle s'écrira $\mathrm{Diag}(1, \dots, 1, 2, \dots, 2)$, avec le nombre de 1 sur la diagonale égal à la dimension de $\mathbb{E}_1$.

Il est très simple également d'écrire les projecteurs associés à la décomposition $\mathbb{E}_1 \oplus \mathbb{E}_2$: sur le sous-espace $\mathbb{E}_2$, l'endomorphisme $u - 2id_{\mathbb{E}}$ est nul, alors qu'il coïncide avec $-id_{\mathbb{E}}$ sur $\mathbb{E}_1$. On a donc évidemment

$$p_1 = 2id_{\mathbb{E}} - u$$

De même $p_2 = u - id_{\mathbb{E}}$ se vérifie aisément (c'est aussi $id_{\mathbb{E}} - p_1$).

L'étude des projecteurs et involutions effectuée à la section **2-2.6** pourrait se mener comme dans l'exemple précédent, pour des endomorphismes ayant respectivement pour polynômes annulateurs

$$X^2 - X = X(X-1) \text{ et } X^2 - 1 = (X-1)(X+1).$$

5-4 Valeurs propres, vecteurs propres

Si u est un endomorphisme d'un espace vectoriel $\mathbb{E}$, les sous-espaces stables les plus simples que l'on soit amené à considérer sont de dimension 1. Comme en dimension 1, les seuls endomorphismes sont les homothéties, on introduit la notion de **valeur propre**.

5-4.1 Valeur propre, sous-espace propre associé

5-4.1.1 Eléments propres d'un endomorphisme, d'une matrice

DÉFINITION 5-4.1 *Soit $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$. Un scalaire $\lambda \in \mathbb{K}$ est dit* ***valeur propre*** *de u si et seulement s'il existe un vecteur $x \in \mathbb{E}$* ***non nul*** *tel que*

$$u(x) = \lambda x$$

DÉFINITION 5-4.2 *Un vecteur $x \in \mathbb{E}$* ***non nul*** *est dit* ***vecteur propre*** *de u si et seulement s'il existe un scalaire $\lambda \in \mathbb{K}$ avec $u(x) = \lambda x$.*

Comme x est supposé non nul, on a $\lambda x = \mu x \Rightarrow \lambda = \mu$. Il y a donc unicité du scalaire λ vérifiant $u(x) = \lambda x$, qui est alors valeur propre de u, conformément à notre première définition. On dit alors indifféremment que x est vecteur propre associé à la valeur propre λ, ou que λ est valeur propre associée au vecteur propre x. Comme

$$u(x) = \lambda x \Leftrightarrow (u - \lambda id_{\mathbb{E}})(x) = 0_{\mathbb{E}}$$

on a la définition équivalente :

DÉFINITION 5-4.3 *Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ si et seulement si*

$$\ker(u - \lambda id_{\mathbb{E}}) \neq \{0_{\mathbb{E}}\}$$

Ce sous-espace vectoriel de $\mathbb{E}$

$$\mathbb{E}_u(\lambda) = \ker(u - \lambda id_{\mathbb{E}})$$

est appelé ***sous-espace propre de l'endomorphisme*** u ***associé à la valeur propre*** λ.

Ce sous-espace propre, par définition non réduit à $\{0_{\mathbb{E}}\}$, est donc formé de tous les vecteurs propres associés à λ et du vecteur nul.

Dire que $x \neq 0_{\mathbb{E}}$ est vecteur propre de u revient exactement à dire que la droite vectorielle engendrée par x est u-stable. On parle alors de $\text{vect}(x)$ comme **direction propre** de u.

DÉFINITION 5-4.4 *L'ensemble des valeurs propres d'un endomorphisme* u *est appelé* ***spectre***[10] *de l'endomorphisme* u. *On le notera* $\text{sp}\, u$.

EXERCICE 5-4.5 Quels sont les spectres des endomorphismes de $\mathbb{K}[X]$ définis par $P \mapsto XP$, et $P \mapsto P'$? Même question avec $f \mapsto f'$ sur $C^{\infty}(\mathbb{R}, \mathbb{R})$.

Lorsque $A \in \mathcal{M}_n(\mathbb{K})$, on travaillera avec l'endomorphisme de $\mathbb{K}^n$ canoniquement associé :

DÉFINITION 5-4.6 $\lambda \in \mathbb{K}$ *est dit valeur propre de la matrice* $A \in \mathcal{M}_n(\mathbb{K})$ *ssi* $A - \lambda I_n$ *n'est pas inversible. Un vecteur propre de* A *est donc un vecteur colonne* $X \in \mathbb{K}^n$ *non nul tel que*

$$AX = \lambda X$$

L'ensemble des vecteurs colonnes vérifiant cette égalité (donc le noyau de la matrice $A - \lambda I_n$*) est le sous-espace propre de* A *pour la valeur propre* λ.

Il est parfois nécessaire de préciser le corps de base : une matrice $A \in \mathcal{M}_n(\mathbb{R})$ peut aussi être considérée comme matrice à coefficients complexes, et on distinguera ses spectres réels et complexes. On a évidemment

$$\underset{\mathbb{R}}{\text{sp}}\, A \subset \underset{\mathbb{C}}{\text{sp}}\, A \quad \text{et} \quad \underset{\mathbb{C}}{\text{sp}}\, A - \underset{\mathbb{R}}{\text{sp}}\, A \subset \mathbb{C} - \mathbb{R}$$

EXERCICE 5-4.7 Déterminer les spectres réels et complexes de

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Si $\mathbb{E}$ est un $\mathbb{K}$-ev de dimension finie, $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ peut être représenté par sa matrice dans une base $\mathcal{B}$ de $\mathbb{E}$:

$$A = M(u, \mathcal{B})$$

[10] Un élément $\lambda \in \mathbb{K}$ est donc dans le spectre ssi $\ker(u - \lambda id_{\mathbb{E}}) \neq \{0_{\mathbb{E}}\}$, c'est à dire si $u - \lambda id_{\mathbb{E}}$ est **non injective**. C'est la terminologie que nous emploierons ultérieurement, contrairement à celle utilisée couramment, pour laquelle le spectre est $\{\lambda \in \mathbb{K} \mid u - \lambda id_{\mathbb{E}} \notin \mathcal{GL}_{\mathbb{K}}(\mathbb{E})\}$. La distinction n'a de toute façon pas lieu d'être si $\mathbb{E}$ est de dimension finie.

Si $x \in \mathbb{E}$ est représenté par la colonne X de ses coordonnées dans $\mathcal{B}$, on a évidemment

$$u(x) = \lambda x \Leftrightarrow AX = \lambda X$$

et par conséquent

$$\operatorname{sp} u = \operatorname*{sp}_{\mathbb{K}} A$$

les éléments des sous-espaces propres de u et A pour une valeur propre λ donnée se correspondant par l'identification entre un vecteur et la colonne de ses coordonnées dans $\mathcal{B}$.

Si $P \in \mathcal{GL}_n(\mathbb{K})$, deux matrices A et $P^{-1}AP$ semblables de $\mathcal{M}_n(\mathbb{K})$ ont même spectre et

$$X \in \mathbb{K}^n \text{ est vecteur propre de } P^{-1}AP \text{ pour } \lambda \Leftrightarrow PX \text{ vecteur propre de } A \text{ pour } \lambda$$

Plus généralement en dimension quelconque, finie ou infinie, on a le théorème

THÉORÈME 5-4.8 Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ et $a \in \mathcal{GL}_{\mathbb{K}}(\mathbb{E})$, les endomorphisme u et aua^{-1} ont même spectre et, pour $\lambda \in \operatorname{sp} u$

$$\ker\left(aua^{-1} - \lambda id_{\mathbb{E}}\right) = a\left(\ker(u - \lambda id_{\mathbb{E}})\right)$$

Démonstration : exercice.

5-4.1.2 Indépendance de sous-espaces propres, spectre de la restriction à un sev stable

Si $(\lambda_i)_{i=1\cdots p}$ sont p scalaires distincts deux à deux, les polynômes $(X - \lambda_i)_{i=1\cdots p}$ sont premiers deux à deux et il découle immédiatement du théorème de décomposition des noyaux le

THÉORÈME 5-4.9 Si $(\lambda_i)_{i=1\cdots p}$ sont p valeurs propres distinctes deux à deux d'un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$, les sous-espaces propres associés sont indépendants.

En particulier, puisqu'une relation de liaison entre les vecteurs d'une famille ne fait intervenir qu'une sous-famille finie extraite, on a le corollaire :

COROLLAIRE 5-4.10 Une famille de vecteurs propres associés à des valeurs propres deux à deux distinctes est une famille libre.

Si $(\lambda_i)_{i=1\cdots p}$ sont p valeurs propres distinctes deux à deux, le sous-espace vectoriel $\mathbb{G} = \bigoplus_{i=1}^{p} \ker(u - \lambda_i id_{\mathbb{E}})$ est alors u-stable. L'endomorphisme induit par u y opère très simplement, comme un recollement d'homothéties[11].

[11]Ceci est encore valable si certains des λ_k ne sont pas dans le spectre de u. Les noyaux correspondants sont alors réduits à $\{0_{\mathbb{E}}\}$

Si $\mathbb{F}$ est un sous-espace de $\mathbb{E}$ **qui est u-stable,** et si $v = u|_{\mathbb{F}}$ est l'endomorphisme de $\mathbb{F}$ induit par u, on a clairement pour $\lambda \in \mathbb{K}$

$$\ker(v - \lambda id_{\mathbb{F}}) = \mathbb{F} \cap \ker(u - \lambda id_{\mathbb{E}})$$

et par conséquent

$$\operatorname{sp} v \subset \operatorname{sp} u$$

REMARQUE 5-4.11 Si $\mathbb{G} = \bigoplus_{i=1}^{p} \ker(u - \lambda_i id_{\mathbb{E}})$ est le sous-espace considéré plus haut, on a de plus

$$\mathbb{G} = \ker \prod_{i=1}^{p} (u - \lambda_i id_{\mathbb{E}})$$

et donc **si $\mathbb{F}$ est u-stable,** on a pour $v = u|_{\mathbb{F}}$

$$\begin{aligned}\mathbb{F} \cap \bigoplus_{i=1}^{p} \ker(u - \lambda_i id_{\mathbb{E}}) &= \ker \prod_{i=1}^{p} (v - \lambda_i id_{\mathbb{F}}) \\ &= \bigoplus_{i=1}^{p} \ker(v - \lambda_i id_{\mathbb{F}}) = \bigoplus_{i=1}^{p} \mathbb{F} \cap \ker(u - \lambda_i id_{\mathbb{E}})\end{aligned}$$

résultat remarquable, si l'on pense qu'en général il n'y a pas distributivité de l'intersection par rapport à la somme directe.

5-4.1.3 Valeurs propres et polynômes

Si $x \in \mathbb{E}$ est vecteur propre de $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ pour une valeur propre λ, on prouve par récurrence sur p que

$$\forall p \in \mathbb{N} \quad u^p(x) = \lambda^p x$$

ce qui donne par linéarité

THÉORÈME 5-4.12 Si x est vecteur propre de u pour la valeur propre λ, pour $P \in \mathbb{K}[X]$ quelconque x est vecteur propre de $P(u)$ pour la valeur propre $P(\lambda)$. De même, lorsque $u \in \mathcal{GL}_{\mathbb{K}}(\mathbb{E})$ (ce qui entraîne que 0 n'est pas valeur propre de u), x est vecteur propre de u^{-1} pour la valeur propre $\frac{1}{\lambda}$.

Démonstration : On a $u(x) = \lambda x \Rightarrow u^{-1}(\lambda x) = x$ soit

$$u^{-1}(x) = \frac{1}{\lambda} x \quad \blacksquare$$

On a donc évidemment

$$P(\operatorname{sp} u) \subset \operatorname{sp} P(u)$$

Cette inclusion peut être stricte : si u est l'endomorphisme de $\mathbb{R}^2$ canoniquement associé à la matrice

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

on a $\operatorname{sp} u = \emptyset$ alors que $u^2 = -id_{\mathbb{E}}$ donc $\operatorname{sp} u^2 = \{-1\}$. Nous verrons cependant ultérieurement que, si u est un endomorphisme d'un $\mathbb{C}$-ev de dimension finie, l'inclusion précédente sera toujours une égalité. On s'en doute, le théorème de D'Alembert est derrière ce résultat.

COROLLAIRE 5-4.13 Si P est un polynôme annulateur pour un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ toute valeur propre de u est racine de P :

$$\operatorname{sp} u \subset \{\alpha \in \mathbb{K} \mid P(\alpha) = 0\}$$

Il en résulte que le spectre d'un endomorphisme qui possède un polynôme minimal (ce qui est toujours le cas en dimension finie) est nécessairement fini.

Si on connaît un polynôme minimal $\pi_u \in \mathbb{K}[X]$, ses racines **dans** $\mathbb{K}$ sont exactement les valeurs propres de u :

THÉORÈME 5-4.14 Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ possède un polynôme minimal $\pi_u \in \mathbb{K}[X]$ on a

$$\operatorname{sp} u = \{\alpha \in \mathbb{K} \mid \pi_u(\alpha) = 0\}$$

Démonstration : On a déjà d'après le corollaire précédent

$$\operatorname{sp} u \subset \{\alpha \in \mathbb{K} \mid \pi_u(\alpha) = 0\}$$

Réciproquement, si $\alpha \in \mathbb{K}$ est une racine de π_u de multiplicité p, on a

$$\pi_u = (X - \alpha)^p R$$

avec $R \in \mathbb{K}[X]$ et $R(\alpha) \neq 0$ et donc $\operatorname{PGCD}((X-\alpha)^p, R) = 1$. Le théorème de décomposition des noyaux donne alors

$$\mathbb{E} = \ker \pi_u(u) = \ker (u - \alpha id_{\mathbb{E}})^p \oplus \ker R(u)$$

Si α n'était pas dans le spectre de u, l'endomorphisme $u - \alpha id_{\mathbb{E}}$ (et donc *a fortiori* $(u - \alpha id_{\mathbb{E}})^p$) serait injectif, et on aurait donc

$$\mathbb{E} = \ker R(u) \text{ et donc } R(u) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{E})}$$

ce qui n'est pas possible, puisque R est un diviseur strict de π_u. ∎

5-4.2 Cas de la dimension finie : polynôme caractéristique

En dimension finie, le déterminant sera un outil théorique pour caractériser les valeurs propres. Il faudra cependant bien garder à l'esprit que la caractérisation géométrique des valeurs propres (λ est valeur propre de u s'il existe une droite vectorielle sur laquelle u opère comme l'homothétie de rapport λ) est souvent plus simple d'utilisation.

5-4.2.1 Polynôme caractéristique d'une matrice carrée

Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}} \in \mathcal{M}_n(\mathbb{K})$. Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de A ssi $A - \lambda I_n$ n'est pas inversible c'est-à-dire ssi

$$\det(A-\lambda I_n) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0$$

Ceci amène à la définition suivante :

DÉFINITION 5-4.15 *Si $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}} \in \mathcal{M}_n(\mathbb{K})$, on appelle **polynôme caractéristique de la matrice** A le polynôme*[12] $\chi_A(X) \in \mathbb{K}[X]$ *défini par*

$$\chi_A(X) = \det(A - XI_n) = \begin{vmatrix} a_{11}-X & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-X & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-X \end{vmatrix}$$

Les racines $\lambda \in \mathbb{K}$ de ce polynôme sont exactement les valeurs propres de A (plus précisément les éléments de $\mathrm{sp}_{\mathbb{K}}(A)$).

THÉORÈME 5-4.16 Le polynôme caractéristique d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est un polynôme de degré exactement n qui peut s'écrire

$$\chi_A(X) = (-1)^n X^n + (-1)^{n-1}\operatorname{trace}(A)X^{n-1} + \cdots + \det A$$

Les matrices A et ${}^tA$ ont mêmes polynômes caractéristiques :

$$\chi_A(X) = \chi_{{}^tA}(X)$$

[12]Pour respecter le programme, nous parlons de déterminant d'une matrice carrée à coefficients dans un corps. Ici, on travaille avec le corps $\mathbb{K}(X)$ des fractions rationnelles à coefficients dans $\mathbb{K}$ (qui contient l'anneau $\mathbb{K}[X]$ des polynômes). Mais comme tous les coefficients de la matrice $B = A - XI_n \in \mathcal{M}_n(\mathbb{K}(X))$ sont en fait des polynômes de $\mathbb{K}[X]$, la formule

$$\det B = \sum_{\sigma\in S_n} \varepsilon(\sigma)\, b_{1\sigma(1)}\cdots b_{n\sigma(n)}$$

montre que ce déterminant est aussi dans $\mathbb{K}[X]$. Si le corps $\mathbb{K}$ est infini (ce qui est usuellement le cas avec $\mathbb{K}$ sous-corps de $\mathbb{C}$), on peut éviter ces contorsions en considérant la **fonction polynôme** $\lambda \mapsto \det(A - \lambda I_n)$.

Démonstration : Avec $B = A - XI_n = (b_{ij}(X))_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}} \in \mathcal{M}_n(\mathbb{K}[X])$, on a

$$\chi_A(X) = \sum_{\sigma\in S_n} \varepsilon(\sigma)\, b_{1\sigma(1)}(X)\cdots b_{n\sigma(n)}(X)$$

est somme de $n!$ polynômes produits de n polynômes de degrés inférieurs ou égaux à 1. Il est donc clair que le degré de $\chi_A(X)$ est inférieur ou égal à n. De plus, dans les $n!$ polynômes considérés plus haut, il en est un seul de degré exactement n, celui pour lequel les n polynômes $b_{i\sigma(i)}(X)$ sont exactement de degré 1, c'est-à-dire le terme correspondant à $\sigma = id_{\{1\ldots n\}}$. Pour tous les autres termes, il y a au moins deux indices i pour lesquels $\sigma(i) \neq i$, et ces termes sont donc de degré inférieur ou égal à $n-2$. On a donc bien prouvé que χ_A est de degré n et que ses termes de degré $\geqslant n-1$ proviennent du produit

$$\prod_{i=1}^{n} (a_{ii} - X)$$

et valent donc bien $(-1)^n X^n + (-1)^{n-1}\,\mathrm{trace}(A)X^{n-1}$. Enfin, le terme constant de χ_A est

$$\chi_A(0) = \det(A - 0I_n) = \det A \quad \blacksquare$$

COROLLAIRE 5-4.17 Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ a au plus n valeurs propres dans $\mathbb{K}$ (ou dans tout sur-corps $\mathbb{L}$ de $\mathbb{K}$).

THÉORÈME 5-4.18 Le polynôme caractéristique est un invariant de similitude.

Démonstration : Si $A \in \mathcal{M}_n(\mathbb{K})$ et $P \in \mathcal{GL}_n(\mathbb{K})$, on a

$$P^{-1}AP - XI_n = P^{-1}(A - XI_n)P$$

relation de similitude dans $\mathcal{M}_n(\mathbb{K}(X))$. La propriété en découle puisque le déterminant est un invariant de similitude. $\blacksquare$

Comme un polynôme formel est déterminé par ses coefficients, l'énoncé qui précède signifie que tous les coefficients du polynôme caractéristique sont des invariants de similitude. On retrouve ainsi cette propriété pour la trace.

EXERCICE 5-4.19 Si $A \in \mathcal{GL}_n(\mathbb{K})$, comparer les polynômes caractéristiques des matrices A et A^{-1}.

EXERCICE 5-4.20 Exprimer le coefficient du terme de degré 1 dans $\chi_A(X)$ à l'aide de la trace de la comatrice de A. La quantité ainsi obtenue est donc aussi un invariant de similitude. (Indication : on pourra utiliser **4-4.4.5**).

Qui dit invariant de similitude dit quantité attachée autant aux endomorphismes qu'aux matrices. La définition qui suit a donc un sens :

DÉFINITION 5-4.21 *Si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel $\mathbb{E}$ de dimension n, on appelle* ***polynôme caractéristique de l'endomorphisme*** *u le polynôme caractéristique de la matrice représentative de u dans une base quelconque de $\mathbb{E}$. On notera χ_u ce polynôme.*

On a donc toujours

$$\chi_u(X) = (-1)^n X^n + (-1)^{n-1} \operatorname{trace}(u) X^{n-1} + \cdots + \det u$$

De plus si $A = M(u, B)$ et $\lambda \in \mathbb{K}$,

$$\det(u - \lambda id_{\mathbb{E}}) = \det(A - \lambda I_n) = \chi_A(\lambda) = \chi_u(\lambda)$$

et par conséquent, **les racines dans $\mathbb{K}$ du polynôme χ_u sont exactement les valeurs propres de l'endomorphime u.** Le corollaire **5-4**.17 est donc valable si on remplace "matrice de $\mathcal{M}_n(\mathbb{K})$" par "endomorphisme d'un $\mathbb{K}$-ev de dimension n" (avec la précision que les valeurs propres sont par définition dans $\mathbb{K}$, il ne peut y avoir ici d'ambiguïté sur le corps de base).

Le théorème de D'Alembert donne immédiatement :

THÉORÈME 5-4.22 Un endomorphisme d'un espace complexe de dimension finie possède au moins une valeur propre.

On notera que ce résultat n'est plus valable en dimension infinie ou sur un espace réel (des contre-exemples ont déjà été donnés).

5-4.2.2 Multiplicité d'une valeur propre

Les définitions seront données dans le cas d'un endomorphisme, elles sont identiques dans le cas des matrices.

DÉFINITION 5-4.23 *Si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel et $\lambda \in \mathbb{K}$ est une valeur propre de u, on appelle* ***multiplicité de la valeur propre*** *λ sa multiplicité comme racine du polynôme caractéristique de u. On la notera*[13] $\operatorname{mul}_u(\lambda) \in \mathbb{N}^*$:

$$p = \operatorname*{mul}_u(\lambda) \Leftrightarrow \chi_u(X) = (X - \lambda)^p R(X) \;\; \textit{avec } R(\lambda) \neq 0$$

Un endomorphisme d'un espace de dimension n possède donc au plus n valeurs propres, chacune d'entre elles étant comptée autant de fois que son ordre de multiplicité.

Un calcul évident donne :

PROPOSITION 5-4.24 Les valeurs propres d'une matrice triangulaire supérieure sont ses coefficients diagonaux, la multiplicité d'une valeur propre étant le nombre de fois qu'elle apparaît sur cette diagonale.

[13] Si $\lambda \notin \operatorname{sp} u$, on pourra poser par convention $\operatorname{mul}_u(\lambda) = 0$.

Le polynôme caractéristique d'une matrice triangulaire supérieure est donc scindé.

Lorsque le polynôme caractéristique d'un endomorphisme u est **scindé** (ce qui est bien sûr toujours le cas si $\mathbb{K} = \mathbb{C}$), les relations entre coefficients et racines donnent :

THÉORÈME 5-4.25 Si u est un endomorphisme d'un espace de dimension n dont le polynôme caractéristique est scindé, si $(\lambda_i)_{1 \leqslant i \leqslant n}$ est la liste de ses valeurs propres (chacune écrite autant de fois que son ordre de multiplicité) on a

$$\operatorname{trace} u = \sum_{i=1}^{n} \lambda_i \quad \textbf{et} \quad \det u = \prod_{i=1}^{n} \lambda_i$$

On a en particulier un résultat analogue pour une matrice $A \in \mathcal{M}_n(\mathbb{C})$ (ou de $\mathcal{M}_n(\mathbb{R})$, en travaillant avec le spectre complexe $\operatorname{sp}_{\mathbb{C}} A$).

Une matrice et sa transposée ayant même polynôme caractéristique ont évidemment mêmes valeurs propres avec les mêmes multiplicités. Nous verrons plus loin (*cf.* section **5-4.2.5**) une interprétation géométrique des vecteurs propres d'une transposée.

5-4.2.3 Polynôme caractéristique d'une restriction à un sous-espace stable

THÉORÈME 5-4.26 Si u est un endomorphisme d'un $\mathbb{K}$-ev $\mathbb{E}$ de dimension finie et si $\mathbb{F} \subset \mathbb{E}$ est un sous-espace stable par u, le polynôme caractéristique $\chi_v(X)$ de l'endomorphisme $v = u|_{\mathbb{F}}$ induit sur $\mathbb{F}$ est un diviseur du polynôme caractéristique de u.

Démonstration : Il suffit pour calculer $\chi_u(X)$ de travailler dans une base $\mathcal{B} = \mathcal{B}_{\mathbb{F}} \cup \mathcal{B}_{\mathbb{S}}$ adaptée à $\mathbb{F}$. Dans cette base, on a

$$M = \mathcal{M}(u, \mathcal{B}) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

Donc, si $\mathbf{E}$ est de dimension n et $\mathbb{F}$ de dimension p

$$\det(M - XI_n) = \begin{vmatrix} A - XI_p & B \\ 0 & C - XI_{n-p} \end{vmatrix}$$

et

$$\chi_u(X) = \det(A - XI_p)\det(C - XI_{n-p}) = \chi_v(X)\det(C - XI_{n-p})$$

ce qui prouve le résultat. ■

On en déduit le corollaire important donnant, à la lecture du polynôme caractéristique, une information (incomplète) sur la dimension d'un sous-espace propre.

COROLLAIRE 5-4.27 Si λ est valeur propre d'un endomorphisme en dimension finie

$$1 \leqslant \dim \mathbb{E}_u(\lambda) \leqslant \operatorname{mul}_u(\lambda)$$

Démonstration : Si $m = \dim \mathbb{E}_u(\lambda)$, le polynôme caractéristique de l'endomorphisme induit par u sur $\mathbb{E}_u(\lambda)$ (qui n'est autre, par définition que $\lambda id_{\mathbb{E}_u(\lambda)}$) vaut $(\lambda - X)^m$ et divise $\chi_u(X)$. On a donc bien $m \leqslant \operatorname{mul}_u(\lambda)$. ∎

Il est à remarquer qu'on ne peut dire plus *a priori* : si $u \in \mathcal{L}(\mathbf{R}^n)$ a pour matrice dans la base canonique

$$A = \begin{pmatrix} \begin{array}{|cccc|} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & \ddots & \ddots & 1 \\ 0 & 0 & 0 & 1 \end{array} & 0 \\ 0 & I_{n-p} \end{pmatrix}$$

où $1 \leqslant p \leqslant n$ est quelconque et où le premier bloc de A a la taille p, la seule valeur propre de A est 1 avec la multiplicité n, alors que le théorème du rang montre facilement que le sous-espace propre correspondant $\ker(A - I_n)$ a pour dimension $n - (p - 1)$ qui peut prendre toute valeur entre 1 et n.

En tout cas, **le sous-espace propre associé à une valeur propre simple (de multiplicité 1) est nécessairement une droite vectorielle.**

Enfin, le résultat sur le déterminant d'une matrice carrée diagonale par blocs donne immédiatement

THÉORÈME 5-4.28 Si $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_j \oplus \cdots \oplus \mathbb{E}_m$ est somme directe de sous-espace u-stables et si on note

$$u_j = u|_{\mathbb{E}_j} \in \mathcal{L}_{\mathbf{K}}(\mathbb{E}_j)$$

l'endomorphisme de $\mathbb{E}_j$ induit par u, on a

$$\chi_u(X) = \prod_{j=1}^{m} \chi_{u_j}(X)$$

5-4.2.4 Théorème de Cayley-Hamilton

Le résultat de cette section est un peu miraculeux et illustre le lien entre polynôme caractéristique et polynôme minimal d'un endomorphisme d'un $\mathbb{K}$-ev de dimension finie. On sait déjà que les racines (dans $\mathbb{K}$) de ces deux polynômes sont exactement les valeurs propres de l'endomorphisme (*cf.* section **5-4.2.1** et théorème **5-4.14**).

THÉORÈME 5-4.29 Si u est un endomorphisme d'un espace $\mathbb{E}$ de dimension finie

$$\chi_u(u) = 0_{\mathcal{L}_{\mathbf{K}}(\mathbb{E})}$$

De manière équivalente, si $A \in \mathcal{M}_n(\mathbb{K})$

$$\chi_A(A) = 0_n$$

Démonstration : Nous donnerons une démonstration matricielle de ce théorème, qui a le mérite d'être rapide, mais qui est aussi plus obscure qu'une démonstration "géométrique", qui ferait intervenir la notion de valeur propre, notion que nous n'utiliserons pas ici. Supposons d'abord le corps $\mathbb{K}$ **infini** (ce qui nous permet d'utiliser concrètement des fonctions polynômes plutôt que de manipuler des polynômes formels). Si A est une matrice de $\mathcal{M}_n(\mathbb{K})$, on considère l'application

$$\mathbb{K} \to \mathcal{M}_n(\mathbb{K}) \quad s \mapsto {}^t[\text{comatrice}(A - sI_n)].$$

Il est clair que les coefficients de la matrice ${}^t[\text{comatrice}(A - sI_n)]$ sont des polynômes en s de degré inférieur ou égaux à $n-1$ (dont les coefficients sont définis sans ambiguïté, puisque $\mathbb{K}$ est infini). On peut donc écrire, en travaillant coefficient par coefficient

$$\forall s \in \mathbb{K} \qquad {}^t[\text{comatrice}(A - sI_n)] = A_0 + sA_1 + \cdots + s^{n-1}A_{n-1}$$

où les matrices $A_i \in \mathcal{M}_n(\mathbb{K})$ ne dépendent pas de s et sont uniques. Soit

$$\chi_A(s) = \alpha_0 + \alpha_1 s + \cdots + \alpha_n s^n$$

le polynôme caractéristique de A (on sait que $\alpha_0 = \det A$, $\alpha_n = (-1)^n$, mais peu importe...). On a

$$\forall s \in \mathbb{K} \quad (A - sI_n)\,{}^t[\text{comatrice}(A - sI_n)] = \chi_A(s)\,I_n$$

(*cf.* section **4-4**.22). Pour $s \in \mathbb{K}$ arbitraire, on obtient en développant le produit de gauche

$$AA_0 + s(AA_1 - A_0) + s^2(AA_2 - A_1) + \cdots \\ + s^{n-1}(AA_{n-1} - A_{n-2}) - s^n A_{n-1} = \chi_A(s)\,I_n$$

Travaillant coefficient par coefficient et puisque $\mathbb{K}$ est infini, on obtient par identification

$$\begin{cases} AA_0 = \alpha_0 I_n \\ AA_1 - A_0 = \alpha_1 I_n \\ \vdots \\ AA_i - A_{i-1} = \alpha_i I_n \\ \vdots \\ AA_{n-1} - A_{n-2} = \alpha_{n-1} I_n \\ -A_{n-1} = \alpha_n I_n \end{cases}$$

Multipliant (à gauche) la première égalité par I_n, la seconde par A, la suivante par A^2, etc... la dernière par A^n et faisant la somme des égalités obtenues, on obtient bien après simplifications

$$\chi_A(A) = 0_n$$

Si le corps $\mathbb{K}$ est fini, on peut remarquer que les calculs précédents peuvent être menés dans tout corps infini $\mathbb{L}$ contenant $\mathbb{K}$, par exemple le corps $\mathbb{K}(Y)$ des fractions rationnelles à une indéterminée Y à coefficients dans $\mathbb{K}$. La démonstration est alors identique, même si les calculs intermédiaires se font *a priori* dans $\mathcal{M}_n(\mathbb{L})$. ■

COROLLAIRE 5-4.30 Le polynôme caractéristique d'un endomorphisme en dimension finie (ou d'une matrice carrée) est un multiple de son polynôme minimal. Ce polynôme minimal est donc, en dimension n, de degré inférieur ou égal à n. En particulier (cf. théorème 5-2.7)

$$\dim \mathbf{E} = n \Rightarrow \forall u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}) \quad \dim \mathbb{K}[u] \leqslant n$$

5-4.2.5 Exercice : hyperplan stable par un endomorphisme, une matrice

Soit $\mathcal{B}$ une base d'un $\mathbb{K}$-ev de dimension n, et $A = M(u, \mathcal{B}) \in \mathcal{M}_n(\mathbb{K})$ la matrice représentative de $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ dans la base $\mathcal{B}$. On cherche à caractériser les **hyperplans u-stables** : un hyperplan $\mathbb{H}$ est représenté comme noyau d'une forme linéaire non nulle ϕ, dont la matrice dans la base $\mathcal{B}$ est une ligne

$$L = (\alpha_1, \dots, \alpha_n) = M(\phi, \mathcal{B}, \{1_{\mathbb{K}}\})$$

(Ces coefficients représentent également les coordonnées de ϕ dans la base duale $\mathcal{B}^*$). Montrer que $\mathbb{H}$ est u-stable ssi la ligne $L \neq 0$ vérifie

$$\exists \mu \in \mathbb{K} \qquad LA = \mu A$$

(ce qui peut s'écrire également ${}^tA{}^tL = \mu{}^tL$: **la recherche des hyperplans stables revient en fait à la recherche des vecteurs propres de la matrice transposée de A**).

5-5 Diagonalisation

Nous nous placerons dans cette section dans le cadre d'un espace vectoriel $\mathbb{E}$ **de dimension finie $n > 0$**.

5-5.1 Endomorphismes et matrices diagonalisables

5-5.1.1 Définition.

Soit $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$. Si le spectre de u n'est pas vide, il est fini (de cardinal inférieur à la dimension de l'espace) et si $(\lambda_i)_{1 \leqslant i \leqslant p}$ sont les valeurs propres **distinctes** de u, on sait que

$$\mathbb{G} = \bigoplus_{i=1}^{p} \ker(u - \lambda_i id_{\mathbb{E}}) = \bigoplus_{i=1}^{p} \mathbb{E}_u(\lambda_i)$$

est un sous-espace de $\mathbb{E}$ stable par u, sur lequel u opère comme recollement d'homothéties. Le cas le plus favorable est celui où $\mathbb{G} = \mathbb{E}$, où la description géométrique de u est particulièrement simple **sur la totalité de l'espace $\mathbb{E}$** :

DÉFINITION 5-5.1 *Un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ est* ***diagonalisable*** *si son spectre n'est pas vide et si l'espace* $\mathbf{E}$ *est somme directe des sous-espaces propres de u.*

Comme des sous-espaces propres associés à des valeurs propres distinctes deux à deux sont indépendants, il est équivalent de dire que la somme des dimensions des sous-espaces propres de u est égale à la dimension de $\mathbf{E}$.

DÉFINITION 5-5.2 *Si u est* ***diagonalisable*** *et $(\lambda_i)_{1\leqslant i\leqslant p}$ sont les valeurs propres* ***distinctes*** *de u, on appelle* ***projecteurs spectraux*** *associés à u la famille de projecteurs associés à la décomposition de l'espace*

$$\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{E}_u(\lambda_i)$$

On notera p_{λ_i} le projecteur sur $\mathbb{E}_u(\lambda_i)$ parallèlement à $\bigoplus_{j\neq i} \mathbb{E}_u(\lambda_j)$.

Si u est diagonalisable et si $x \in \mathbb{E}$ se décompose dans la somme directe des sous-espaces propres sous la forme

$$x = x_1 + \cdots + x_p$$

on aura

$$u(x) = \lambda_1 x_1 + \cdots + \lambda_p x_p$$

ce qui montre que u se décompose comme combinaison linéaire des projecteurs spectraux

$$u = \sum_{i=1}^{p} \lambda_i p_{\lambda_i} = \sum_{\lambda \in \mathrm{sp}\, u} \lambda p_\lambda$$

EXERCICE 5-5.3 Montrer qu'alors pour tout polynôme $P \in \mathbb{K}[X]$ on a

$$P(u) = \sum_{\lambda \in \mathrm{sp}\, u} P(\lambda)\, p_\lambda$$

Effectuer des calculs polynomiaux sur un endomorphisme diagonalisable revient en quelque sorte à faire ces mêmes calculs sur les valeurs propres de u.

EXERCICE 5-5.4 Montrer que réciproquement, si $\mathbb{E} = \mathbb{E}_1 \oplus \cdots \oplus \mathbb{E}_j \oplus \cdots \oplus \mathbb{E}_m$ est une décomposition de l'espace en sous-espaces non nuls, si $(p_i)_{1\leqslant i\leqslant m}$ est la famille de projecteurs associée à cette décomposition et $(\alpha_i)_{1\leqslant i\leqslant m}$ est une famille de scalaires quelconques, l'endomorphisme

$$v = \sum_{i=1}^{m} \alpha_i p_i$$

(c'est-à-dire tout endomorphisme obtenu en recollant des homothéties) est diagonalisable. Quel est son spectre ? Déterminer ses sous-espaces propres (attention : les $(\alpha_i)_{1\leqslant i\leqslant m}$ n'ont pas été supposés distincts deux à deux).

5-5.1.2 Caractérisation

THÉORÈME 5-5.5 Si u est un endomorphisme d'un $\mathbb{K}$-ev $\mathbb{E}$ de dimension finie n, les propriétés suivantes sont équivalentes :

i) **u est diagonalisable**
ii) **Il existe une base de E composée de vecteurs propres de u**
iii) **Il existe une base de $\mathbb{E}$ dans laquelle la matrice de u est diagonale**
iv) $\chi_u(X)$ **est scindé dans** $\mathbb{K}[X]$ **et** $\forall \lambda \in \mathrm{sp}\, u \quad \dim \mathbb{E}_u(\lambda) = \mathrm{mul}_u(\lambda)$

Démonstration : Les implications $i) \Rightarrow ii)$ et $ii) \Rightarrow iii)$ sont évidentes. Supposons que, dans une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ la matrice de u soit $A = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n)$, le calcul du déterminant dans cette base donne

$$\chi_u(X) = \prod_{i=1}^{n} (\alpha_i - X)$$

Le polynôme $\chi_u(X)$ est donc scindé, et $\alpha_1, \ldots, \alpha_n$ est la liste des valeurs propres de u, chacune d'entre elles étant écrite autant de fois que sa multiplicité. Si $\lambda \in \mathrm{sp}\, u$ et si on note $I_\lambda = \{i \in [1..n] \mid \alpha_i = \lambda\}$, on a donc

$$\underset{u}{\mathrm{mul}}(\lambda) = \mathrm{card}\, I_\lambda$$

Comme de manière évidente $\mathrm{vect}\,(e_i)_{i \in I_\lambda} \subset \mathbb{E}_u(\lambda)$, on obtient

$$\underset{u}{\mathrm{mul}}(\lambda) \leqslant \dim \mathbb{E}_u(\lambda)$$

L'inégalité inverse étant toujours vraie, on a bien $iii) \Rightarrow iv)$. Enfin, si on suppose $iv)$, comme $d^\circ \chi_u = n$, on a

$$\sum_{\lambda \in \mathrm{sp}\, u} \dim \mathbb{E}_u(\lambda) = \sum_{\lambda \in \mathrm{sp}\, u} \underset{u}{\mathrm{mul}}(\lambda) = d^\circ \chi_u = n$$

ce qui entraîne que u est diagonalisable. ■

REMARQUE 5-5.6 La propriété *iii)* précédente aurait pu s'exprimer sous la forme :
"L'espace est somme directe de sous-espaces sur lesquels u opère comme une homothétie".

COROLLAIRE 5-5.7 (cas particulier important) Un endomorphisme dont le polynôme caractéristique est scindé et dont toutes les racines sont simples est diagonalisable et ses sous-espaces propres sont des droites vectorielles.

Démonstration : La caractérisation $iv)$ est immédiatement vérifiée puisque, pour $\lambda \in \mathrm{sp}\, u$

$$1 \leqslant \dim \mathbb{E}_u(\lambda) \leqslant \underset{u}{\mathrm{mul}}(\lambda) = 1$$

donc toutes ces inégalités sont des égalités. **Attention ! Il s'agit ici d'une condition suffisante pour que u soit diagonalisable. Elle n'est évidemment pas nécessaire comme le montre l'exemple $u = id_{\mathbb{E}}$!** ■

Comme tout polynôme d'une matrice diagonale est une matrice diagonale, et puisque l'inverse d'une matrice diagonale de $\mathcal{GL}_n(\mathbb{K})$ est aussi diagonale, on a

COROLLAIRE 5-5.8 Si u est diagonalisable, tout élément de $\mathbb{K}[u]$ est diagonalisable. Si u est de plus inversible, u^{-1} est aussi diagonalisable.

5-5.1.3 Cas des matrices carrées

Les définitions données sur les endomorphismes peuvent se transposer aux matrices :

DÉFINITION 5-5.9 *$A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable si et seulement si l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé est diagonalisable. D'après la caractérisation iii) du théorème* **5-5**.*5, cela équivaut clairement à dire que A* ***est semblable dans l'algèbre $\mathcal{M}_n(\mathbb{K})$ à une matrice diagonale :***

$$\exists\, P \in \mathcal{GL}_n(\mathbb{K}) \quad D = P^{-1}AP \text{ est diagonale.}$$

Les caractérisations du théorème **5-5**.5 sont valables pour les matrices carrées, avec les abus de langage usuels consistant à identifier une matrice et l'endomorphisme de $\mathbb{K}^n$ canoniquement associé. La caractérisation *iii*) montre encore qu'un endomorphisme est diagonalisable ssi sa matrice dans une base quelconque est diagonalisable.

Attention ! La définition précédente a été donnée **en travaillant exclusivement sur $\mathbb{K}$.** Il peut y avoir ambiguïté lorsqu'on travaille avec une extension $\mathbb{L}$ de $\mathbb{K}$ (le cas classique étant $\mathbb{K} = \mathbb{R}$ et $\mathbb{L} = \mathbb{C}$). On distinguera pour $A \in \mathcal{M}_n(\mathbb{K})$ "être $\mathbb{K}$-diagonalisable" et "être $\mathbb{L}$-diagonalisable". Par exemple, il est facile de vérifier que

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

est $\mathbb{C}$-diagonalisable mais n'est pas $\mathbb{R}$-diagonalisable.

5-5.1.4 Pratique de la réduction

Soit u un endomorphisme diagonalisable. **Diagonaliser** u, c'est déterminer son spectre et, pour chaque valeur propre, déterminer le sous-espace propre correspondant, en général en en donnant une base. La réunion de ces bases est alors une base de $\mathbb{E}$ (ce qui caractérise le fait que u soit diagonalisable !) formée de vecteurs propres de u, dans laquelle la matrice de u est diagonale.
De même **diagonaliser une matrice carrée** A c'est trouver une matrice P inversible avec $P^{-1}AP$ diagonale.
Comme on connaît souvent un endomorphisme par sa matrice dans une base, les deux notions sont liées. Pour diagonaliser une matrice $A \in \mathcal{M}_n(\mathbb{K})$:

- On commence par déterminer le spectre de A (attention, pas toujours en utilisant le polynôme caractéristique. Cependant, si on connaît $\chi_A(X)$, inutile d'aller plus loin si celui-ci n'est pas scindé).

- Pour chaque valeur propre $\lambda \in \operatorname{sp} A$, recherche d'une base du sous-espace propre associé, c'est-à-dire une base de l'espace des vecteurs colonnes X solutions de

$$(A - \lambda I_n) X = 0$$

Si on connaît $\operatorname{mul}_A(\lambda)$, ne pas espérer diagonaliser A si cette base contient moins de $\operatorname{mul}_A(\lambda)$ vecteurs, c'est-à-dire (théorème du rang) si le rang de $A - \lambda I_n$ n'est pas égal à $n - \operatorname{mul}_A(\lambda)$

- Enfin, si on obtient par les calculs précédents, et en regroupant les bases obtenues, n vecteurs propres (forcément indépendants si on n'a pas fait d'erreur de calcul...) $X_1, X_2, \ldots, X_n$ pour les valeurs propres (non nécessairement distinctes) respectives $\lambda_1, \ldots, \lambda_n$, la matrice des vecteurs colonnes $X_1, X_2, \ldots, X_n$

$$P = (X_1 | X_2 | \cdots | X_n)$$

est matrice de passage de la base canonique de $\mathbb{K}^n$ à un base de vecteurs propres de l'endomorphisme canoniquement associé à A, et par conséquent

$$P^{-1} A P = \operatorname{Diag}(\lambda_1, \ldots, \lambda_n) = D$$

EXEMPLE 5-5.10 Décrire géométriquement l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique vaut

$$A = \begin{pmatrix} -9 & -10 & -20 \\ 4 & 5 & 8 \\ 2 & 2 & 5 \end{pmatrix}.$$

On calcule $\chi_A(X) = -(X+1)(X-1)^2$: dans la matrice $A - XI_3$, on remarque que

$$C_1 \leftarrow C_1 + C_2 - C_3$$

permet d'obtenir une première colonne où $1 - X$ se met en facteur. On a donc deux valeurs propres 1 et -1 de multiplicités respectives 2 et 1. A sera diagonalisable ssi le sous-espace propre $E_A(1)$ est de dimension 2 c'est-à-dire si $A - I_n$ est de rang 1, ce qui est bien le cas puisque

$$A - I_n = \begin{pmatrix} -10 & -10 & -20 \\ 4 & 4 & 8 \\ 2 & 2 & 4 \end{pmatrix}$$

a ses vecteurs lignes proportionnels à $(1, 1, 2)$. Le sous-espace $E_A(1)$ est donc le plan de $\mathbb{R}^3$ d'équation

$$x + y + 2z = 0$$

dont on peut choisir une base, par exemple :

$$X_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \text{ et } X_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Le sous-espace $E_A(-1)$ est, on le sait, une droite vectorielle déterminée par l'équation

$$(A + I_n)X = 0$$

équivalente au système (de rang 2, on le sait)

$$\begin{cases} 4x + 6y + 8z = 0 \\ 2x + 2y + 6z = 0 \end{cases}$$

Un vecteur directeur de cette droite peut être obtenu par le "produit vectoriel canonique" des vecteurs $(2,3,4)$ et $(1,1,3)$, respectivement "normaux" aux deux plans dont l'intersection est $E_A(-1)$ (*cf.* remarque à la fin de la section **4-5.2**), soit

$$X_3 = \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}$$

La matrice

$$P = \begin{pmatrix} 1 & 1 & 5 \\ 1 & -1 & -2 \\ -1 & 0 & -1 \end{pmatrix}$$

vérifie donc

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Il est facile de décrire maintenant géométriquement l'endomorphisme associé à la matrice A.[14]

REMARQUE 5-5.11 Pour déterminer le spectre d'une matrice (ou d'un endomorphisme) le calcul du polynôme caractéristique ne s'impose pas toujours. Prenons l'exemple de la matrice

$$A = \begin{pmatrix} a & 1 & 1 & \cdots & 1 \\ 1 & a & 1 & \ddots & \vdots \\ 1 & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & 1 & a & 1 \\ 1 & \cdots & 1 & 1 & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$$

Il n'est pas très difficile, par des manipulations élémentaires, de calculer le polynôme caractéristique de A. Le raisonnement suivant est plus instructif, et permet de diagonaliser très rapidement A (ou plutôt de prouver qu'elle est diagonalisable

[14]Dès le départ, le calcul de A^2 et de la trace de A aurait permis de voir que A est matrice d'une symétrie par rapport à un plan

et de la diagonaliser[15]) : on remarque d'abord qu'en retranchant $a-1$ à tous les coefficients diagonaux de A, on obtient une matrice de rang 1, dont le noyau est l'hyperplan $\mathbb{H}$ de $\mathbb{R}^n$ d'équation

$$\sum_{i=1}^{n} x_i = 0$$

On en déduit que A possède la valeur propre $a-1$, avec un sous-espace propre associé de dimension $n-1$, donc de multiplicité au moins $n-1$. Le polynôme caractéristique de A est donc scindé, avec au moins $n-1$ fois la racine $a-1$, la dernière racine λ étant calculable à l'aide de la trace de A :

$$\lambda + (n-1)(a-1) = \operatorname{trace} A = na$$

ce qui donne

$$\lambda = a + n - 1 \neq a - 1$$

ce qui montre que λ est valeur propre simple de A, le sous-espace propre associé étant une droite vectorielle. **A est diagonalisable.** De plus, il est clair que le vecteur colonne dont toutes les coordonnées valent 1 est vecteur propre pour λ. Il reste à écrire une base de $\mathbb{H}$ pour trouver la matrice de passage, par exemple :

$$P = \begin{pmatrix} 1 & -1 & -1 & \cdots & \cdots & -1 \\ 1 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & 1 & 0 & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 & \vdots \\ 1 & \vdots & \vdots & 0 & \ddots & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

qui vérifie

$$P^{-1}AP = \operatorname{Diag}(a+n-1, a-1, a-1, \ldots, a-1)$$

5-5.1.5 Caractérisation à l'aide de polynômes annulateurs

Les caractérisations des endomorphismes diagonalisables données jusqu'à présent étaient plutôt géométriques. L'utilisation d'un critère plus algébrique est souvent rapide et commode :

THÉORÈME 5-5.12 Un endomorphisme $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ est diagonalisable si et seulement s'il annule un polynôme de $\mathbb{K}[X]$ scindé à racines simples.

Démonstration : Soit P un polynôme annulateur scindé à racine simples, qu'on peut supposer normalisé :

$$P = \prod_{i=1}^{m} (X - \alpha_i)$$

[15] Avec des connaissances sur les espaces euclidiens à venir, un argument très simple prouverait que A est $\mathbb{R}$-diagonalisable

Comme les α_i sont distincts deux à deux, le théorème de décomposition des noyaux donne

$$\mathbb{E} = \ker P(u) = \bigoplus_{i=1}^{m} \ker(u - \alpha_i id_{\mathbb{E}})$$

ce qui montre que l'espace est somme directe de sous-espaces sur lesquels u opère comme une homothétie et u est diagonalisable : la matrice de u dans une base adaptée à la décomposition précédente (en oubliant bien sûr les noyaux réduits éventuellement à $\{0_{\mathbb{E}}\}$) est diagonale. On a d'ailleurs

$$\mathrm{sp}\, u \subset \{\alpha_1, \ldots, \alpha_m\}$$

Réciproquement, si u est diagonalisable et $\mathrm{sp}\, u = \{\lambda_1, \ldots, \lambda_p\}$, (chaque valeur propre n'est ici écrite qu'une fois), le polynôme

$$P = \prod_{i=1}^{p} (X - \lambda_i)$$

est annulateur pour u (toujours à cause du théorème de décomposition des noyaux, voir aussi l'exercice **5-5**.3). ■

COROLLAIRE 5-5.13 Le polynôme minimal d'un endomorphisme u diagonalisable est

$$\pi_u(X) = \prod_{\lambda \in \mathrm{sp}\, u} (X - \lambda)$$

(où chaque valeur propre n'est bien entendu écrite qu'une fois).

Démonstration : Ce polynôme est effectivement annulateur, d'après ce qui précède. Donc $\pi_u(X)$ divise $\prod_{\lambda \in \mathrm{sp}\, u} (X - \lambda)$ De plus, comme toute valeur propre de u est racine de $\pi_u(X)$, on a aussi la relation de divisibilité inverse, ce qui prouve le résultat. ■

Pour un endomorphisme u **diagonalisable** on a donc

$$\begin{cases} \pi_u(X) = \prod_{\lambda \in \mathrm{sp}\, u} (X - \lambda) \\ \chi_u(X) = (-1)^{\dim \mathbf{E}} \prod_{\lambda \in \mathrm{sp}\, u} (X - \lambda)^{\mathrm{mul}_u(\lambda)} \end{cases}$$

Le théorème de Cayley-Hamilton est donc une évidence dans ce cas.

EXERCICE 5-5.14 Soit $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$A^3 - 4A^2 + 5A = 0_n$$

Montrer que le rang de A est pair.

EXERCICE 5-5.15 Soit $A \in \mathcal{M}_2(\mathbb{Z})$ telle que

$$\exists p \in \mathbb{N}^* \quad A^p = I_2$$

Montrer que $A^{12} = I_2$. (Indication pour ces deux exercices : considérer d'abord la matrice A comme appartenant à une algèbre de matrices **complexes**).

5-5.2 Applications de la notion de diagonalisation

5-5.2.1 Recherche de sous-espaces stables par un endomorphisme diagonalisable

THÉORÈME 5-5.16 Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ est diagonalisable et $\mathbb{F} \neq \{0_{\mathbb{E}}\}$ est un sous-espace u-stable, l'endomorphisme $v = u|_{\mathbb{F}}$ induit sur $\mathbb{F}$ est également diagonalisable.

Démonstration : Donnons une démonstration purement algébrique, nous verrons ensuite une interprétation géométrique de ce résultat. Si P est un polynôme annulateur scindé à racines simples, on a par définition $P(u) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{E})}$. Par restriction à $\mathbb{F}$, $P(v) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{F})}$, ce qui prouve que v est diagonalisable. ■

L'interprétation géométrique de ce résultat a déjà été vue à la remarque **5-4**.11 : Si l'endomorphisme u est diagonalisable

$$\mathbb{E} = \bigoplus_{\lambda \in \mathrm{sp}\, u} \mathbb{E}_u(\lambda)$$

et si $\mathbb{F}$ est un sous-espace u-stable

$$\mathbb{F} = \bigoplus_{\lambda \in \mathrm{sp}\, u} (\mathbb{E}_u(\lambda) \cap \mathbb{F})$$

et $\mathbb{F}$ est somme directe de sous-espaces sur lesquels v opère comme un homothétie.

EXERCICE 5-5.17 Déterminer les sous-espaces de $\mathbb{R}^3$ stables par la matrice

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

Pour déterminer les plans stables, on pourra utilisera le théorème précédent. Retrouver ensuite le résultat avec la section **5-4.2.5**.

5-5.2.2 Calculs polynomiaux sur un endomorphisme diagonalisable

En utilisant la décomposition à l'aide des projecteurs spectraux

$$u = \sum_{\lambda \in \mathrm{sp}\, u} \lambda p_\lambda$$

il est clair que

$$\forall Q \in \mathbb{K}[X] \qquad Q(u) = \sum_{\lambda \in \mathrm{sp}\, u} Q(\lambda)\, p_\lambda$$

égalité qui traduit simplement le fait que $Q(u)$ opère sur chaque sous-espace propre $\mathbb{E}_u(\lambda)$ comme l'homothétie de rapport $Q(\lambda)$.

La traduction matricielle de ce résultat est simplement

$$A = P\,\mathrm{Diag}\,(\alpha_1,\dots,\alpha_n)\,P^{-1} \implies$$
$$\forall Q \in \mathbb{K}[X] \quad Q(A) = P\,\mathrm{Diag}\,(Q(\alpha_1),\dots,Q(\alpha_n))\,P^{-1}$$

En particulier pour $p \in \mathbb{N}^*$ ($p \in \mathbb{Z}$ si aucun des α_i n'est nul)

$$A^p = P\,\mathrm{Diag}\,(\alpha_1^p,\dots,\alpha_n^p)\,P^{-1}$$

On se rappellera cependant que, lorsqu'on connaît un polynôme annulateur pour une matrice A (qu'elle soit diagonalisable ou pas), les calculs dans $\mathbb{K}[A]$ peuvent se mener après division euclidienne par ce polynôme. Cela évite parfois, dans le cas diagonalisable, une détermination explicite de la matrice de passage P.

EXERCICE 5-5.18 Montrer que, si u est diagonalisable, les projecteurs spectraux sont dans $\mathbb{K}[u]$ (sans utiliser la remarque **5-3.6**) et les expliciter. (Penser aux polynômes interpolateurs de Lagrange).

5-5.2.3 Commutant d'un endomorphisme (d'une matrice) diagonalisable

Commençons par un résultat énoncé plus haut dans un cadre plus général.

PROPOSITION 5-5.19 Si u et v sont deux endomorphismes (non nécessairement diagonalisables, éventuellement en dimension infinie) qui commutent, les sous-espaces propres de l'un sont stables par l'autre.

Démonstration : On peut simplement dire que le sous-espace vectoriel $\mathbb{E}_u(\lambda) = \ker(u - \lambda id_{\mathbb{E}})$, étant noyau d'un polynôme en u qui commute avec v, est v-stable. Plus concrètement

$$u(x) = \lambda x \Rightarrow u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x) \quad \blacksquare$$

Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$, on appelle **commutant de** u et on note $C(u)$ l'ensemble

$$C(u) = \{v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}) \mid uv = vu\}$$

On vérifie aisément que $C(u)$ **est une sous-algèbre de $L_{\mathbb{K}}(\mathbb{E})$ qui contient** $K[u]$.

THÉORÈME 5-5.20 Si u est un endomorphisme diagonalisable d'un $\mathbb{K}$-ev $\mathbb{E}$, un endomorphisme $v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ commute avec u si et seulement si v laisse stable tous les sous-espaces propres de u.

Démonstration : Cette condition est nécessaire d'après la proposition précédente. Elle est suffisante : si $(\lambda_i)_{1\leqslant i\leqslant p}$ est le spectre de u

et si $x = x_1 + \cdots + x_p$ est la décomposition d'un vecteur de $\mathbb{E}$ dans la somme directe $\bigoplus_{i=1}^{p} \mathbb{E}_u(\lambda_i)$, on a

$$vu(x) = v\left(\sum_{i=1}^{p} \lambda_i x_i\right) = \sum_{i=1}^{p} \lambda_i v(x_i)$$

et comme par hypothèse $v(x_i) \in \mathbb{E}_u(\lambda_i)$, on a $\lambda_i v(x_i) = u(v(x_i))$. On obtient finalement

$$vu(x) = \sum_{i=1}^{p} uv(x_i) = uv\left(\sum_{i=1}^{p} x_i\right) = uv(x) \quad \blacksquare$$

EXERCICE 5-5.21 Résoudre dans $\mathcal{M}_3(\mathbb{R})$ l'équation $X^2 = A$ avec

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

(on remarquera qu'une solution commute nécessairement avec A).

EXERCICE 5-5.22 Si $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$ est diagonalisable, montrer que

$$\dim \mathcal{C}(u) = \sum_{\lambda \in \operatorname{sp} u} \operatorname{mul}_u(\lambda)^2$$

(On pourra travailler matriciellement dans une base bien choisie).

EXERCICE 5-5.23 Si u est diagonalisable montrer que

$$\mathcal{C}(\mathcal{C}(u)) = \{v \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}) \mid \forall w \in \mathcal{C}(u) \quad wv = vw\}$$

est une sous-algèbre de $\mathcal{L}_{\mathbb{K}}(\mathbb{E})$, puis qu'elle est égale à $\mathbb{K}[u]$.

EXERCICE 5-5.24 Si u est un endomorphisme d'un espace de dimension n qui possède n valeurs propres simples, le commutant de u est égal à $\mathbb{K}[u]$ et est de dimension n.

Terminons enfin par un résultat de **diagonalisation simultanée** qui intervient dans de nombreux exercices.

THÉORÈME 5-5.25 Si u et v sont deux endomorphismes diagonalisables d'un espace $\mathbb{E}$, ils commutent si et seulement s'il existe une base de $\mathbb{E}$ formée de vecteurs propres à la fois pour u et v. On dit alors que u et v se diagonalisent dans la même base, ou sont simultanément diagonalisables.

On aura bien sûr un résultat analogue pour les matrices : deux matrices A et B **diagonalisables** de $\mathcal{M}_n(\mathbb{K})$ commutent si et seulement s'il existe une même matrice de passage P telle que $P^{-1}AP$ et $P^{-1}BP$ soient toutes deux diagonales.

Démonstration : La condition est évidemment suffisante : s'il existe une base de vecteurs propres communs à u et v, les matrices de ces endomorphismes dans cette base sont diagonales donc commutent, et en conséquence $uv = vu$. Réciproquement, supposons $uv = vu$ et considérons la décomposition de l'espace

$$\mathbb{E} = \bigoplus_{\lambda \in \mathrm{sp}\, u} \mathbb{E}_u(\lambda)$$

Chacun des sous-espaces $\mathbb{E}_u(\lambda)$ est v-stable, et l'endomorphisme induit sur ce sev par v est diagonalisable (*cf.* théorème **5-5.16**). On peut donc trouver dans chaque $\mathbb{E}_u(\lambda)$ une base formée de vecteurs propres pour v (et évidemment pour u !). La réunion de ces bases diagonalise à la fois u et v. ■

EXERCICE 5-5.26 Généraliser ce résultat à une famille quelconque d'endomorphismes diagonalisables qui commutent deux à deux. Le plus simple techniquement est ici de raisonner par récurrence sur la dimension de l'espace $\mathbb{E}$.

5-6 Trigonalisation

Nous avons vu précédemment des exemples de matrices non diagonalisables. Le but de cette section est de trouver, pour une telle matrice, une similitude qui la transforme en une matrice plus simple. Comme on ne peut espérer une forme diagonale, on essaiera une forme triangulaire supérieure.

5-6.1 Endomorphismes, matrices trigonalisables

DÉFINITION 5-6.1 *Un endomorphisme d'un espace vectoriel* $\mathbf{E}$ *de dimension finie* n *est dit* ***trigonalisable*** *si et seulement s'il existe une base de* $\mathbf{E}$ *où la matrice de* u *est triangulaire supérieure. On dit alors qu'une telle base trigonalise l'endomorphisme* u.

Si $\mathcal{B}$ est une base dans laquelle la matrice de u est triangulaire "supérieure", en inversant l'ordre des vecteurs de $\mathcal{B}$, on obtiendra une base dans laquelle la matrice est triangulaire "inférieure". Remarquons aussi qu'un endomorphisme diagonalisable est *a fortiori* trigonalisable !

Rappelons le résultat déjà énoncé dans le chapitre sur le calcul matriciel :

PROPOSITION 5-6.2 Si u est un endomorphisme d'un espace $\mathbb{E}_n$ de dimension n et si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de $\mathbb{E}_n$, la matrice de u dans la base $\mathcal{B}$ est triangulaire supérieure ssi les sous-espaces du drapeau associé à $\mathcal{B}$ sont tous stables par u :

$$\forall i \qquad u(\mathrm{vect}(e_1, \ldots, e_i)) \subset \mathrm{vect}(e_1, \ldots, e_i)$$

La définition donnée sur les endomorphismes se transpose sur les matrices carrées :

DÉFINITION 5-6.3 *$A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable si et seulement si l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé l'est également. Cela équivaut clairement à dire que A* ***est semblable dans l'algèbre $\mathcal{M}_n(\mathbb{K})$ à une matrice triangulaire supérieure :***

$$\exists P \in \mathcal{GL}_n(\mathbb{K}) \quad T = P^{-1}AP \text{ est triangulaire supérieure.}$$

Il faut ici aussi faire attention si l'on considère une extension $\mathbb{L}$ de $\mathbb{K}$. Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ peut être $\mathbb{L}$-trigonalisable sans être $\mathbb{K}$-trigonalisable.

Comme dans le cas diagonalisable, le caractère trigonalisable peut se traduire en terme de polynôme annulateur :

THÉORÈME 5-6.4 Si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel $\mathbb{E}_n$ de dimension n, les propriétés suivantes sont équivalentes :

i) **u est trigonalisable**
ii) **le polynôme caractéristique χ_u est scindé dans $\mathbb{K}[X]$**
iii) **u possède un polynôme annulateur scindé dans $\mathbb{K}[X]$**

Démonstration : Si u est trigonalisable, on peut calculer $\chi_u(X)$ en travaillant dans une base où la matrice de u est triangulaire. Ce polynôme est donc scindé. On a bien *i)* $\Rightarrow$ *ii)*. L'implication *ii)* $\Rightarrow$ *iii)* est conséquence du théorème de Cayley-Hamilton. Enfin montrons que *iii)* $\Rightarrow$ *i)* en raisonnant par récurrence sur n. Si $n = 1$, il n'y a rien à démontrer. Supposons que tout endomorphisme de tout $\mathbb{K}$-ev de dimension $n-1$ qui annule un polynôme scindé soit trigonalisable. Soit $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$, possédant

$$P(X) = \prod_{i=1}^{m}(X - \alpha_i)$$

comme polynôme annulateur, avec les $\alpha_i \in \mathbb{K}$ (non nécessairement distincts deux à deux). Pour démontrer que u est trigonalisable, **il suffit de trouver un hyperplan $\mathbb{H} \subset \mathbb{E}_n$ stable par u.** En effet, si v est alors l'endomorphisme induit par u sur $\mathbb{H}$, on a

$$P(u) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)} \Rightarrow P(v) = \prod_{i=1}^{m}(v - \alpha_i id_{\mathbb{H}}) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{H})}$$

En appliquant l'hypothèse de récurrence à v, on pourra trouver une base $\mathcal{B}_1$ de $\mathbb{H}$ trigonalisant v. Si on complète cette base (par un vecteur) en une base $\mathcal{B} = \mathcal{B}_1 \cup \{e_n\}$ de $\mathbb{E}_n$, la matrice de u dans $\mathcal{B}$ sera évidemment triangulaire supérieure. Reste donc à prouver l'existence de $\mathbb{H}$.

Comme

$$\prod_{i=1}^{m}(u - \alpha_i id_{\mathbb{E}_n}) = 0_{\mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)}$$

l'un (au moins) des endomorphismes $(u - \alpha_i id_{\mathbb{E}_n})$ est non injectif. Si c'est par exemple le cas avec $i = 1$, $\mathrm{Im}\,(u - \alpha_1 id_{\mathbb{E}_n})$ est un sous-espace strict de $\mathbb{E}_n$ stable par u. Si $\mathbb{H}$ est un hyperplan contenant $\mathrm{Im}\,(u - \alpha_1 id_{\mathbb{E}_n})$, il est évidemment stable par $(u - \alpha_1 id_{\mathbb{E}_n})$ puisque

$$(u - \alpha_1 id_{\mathbb{E}_n})(\mathbb{H}) \subset \mathrm{Im}\,(u - \alpha_1 id_{\mathbb{E}_n}) \subset \mathbb{H}$$

On en déduit facilement que $\mathbb{H}$ est aussi u-stable. ■

COROLLAIRE 5-6.5 Tout endomorphisme d'un espace complexe de dimension finie, toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable.

Par contre, une matice carrée (n, n) à coefficients réels n'est $\mathbb{R}$-trigonalisable que si son polynôme caractéristique possède n racines réelles (comptées avec leurs multiplicités).

5-6.2 Pratique de la trigonalisation

Nous présenterons la technique dans le cas d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, que l'on identifie à l'endomorphisme Φ_A de $\mathbb{K}^n$ canoniquement associé. On suppose que le polynôme caractéristique $\chi_A(X)$ possède n racines (éventuellement multiples) $\alpha_1, \dots, \alpha_n$. (Il est illusoire de vouloir trigonaliser A si on ne peut déterminer ses valeurs propres, puisque ce sont celles-ci qui apparaîtront -forcément- sur la diagonale d'une matrice triangulaire supérieure semblable à A). La démarche suivie ici donnerait d'ailleurs une démonstration plus "algorithmique" de l'équivalence des énoncés i) et ii) du théorème **5-6.4**. Notons $\mathcal{B}_c$ la base canonique de $\mathbb{K}^n$.

Il s'agit ici de construire de proche en proche des vecteurs e_i, $i = 1 \dots n$, d'une base de $\mathbb{K}^n$ avec

$$\forall i \qquad \Phi_A(\mathrm{vect}(e_1, \dots, e_i)) \subset \mathrm{vect}(e_1, \dots, e_i) \qquad (1)$$

Remarquons que e_1 est nécessairement vecteur propre de Φ_A. Cela tombe bien : puisque le polynôme caractéristique est scindé, Φ_A possède au moins une valeur propre, donc au moins une direction propre. On pourrait initialiser le processus de construction en choisissant un vecteur propre arbitraire, mais on peut démarrer un peu plus vite :

- On commence par déblayer le terrain : pour chaque valeur propre de A, on trouve une base du sous-espace propre correspondant. On réunit toute ces bases en un famille libre $\mathcal{F}_p$ de cardinal p. Si $p = n$, A est diagonalisable.

- Sinon, $p < n$. Soit $\mathcal{F}_p = (e_1, \dots, e_p)$ et $\mathbb{E}_p = \mathrm{vect}\,\mathcal{F}_p$. Il est clair que la propriété (1) est vérifiée pour $i = 1 \cdots p$. On complète comme on peut (en général avec des vecteurs de la base canonique bien choisis) $\mathcal{F}_p$ en une base $\mathcal{B}_p = \mathcal{F}_p \cup \mathcal{G}_p$ de $\mathbb{K}^n$, et on détermine la matrice A_p de Φ_A dans cette base :

$$A_p = M(\Phi_A, \mathcal{B}_p) = \begin{pmatrix} T_p & B_p \\ 0 & A'_p \end{pmatrix}$$

avec T_p triangulaire supérieure (en fait diagonale, car les vecteurs de $\mathcal{F}_p$ sont propres pour Φ_A, mais cela n'importe pas dans la suite du raisonnement).

- Notons $\mathcal{G}_p = (u_1, \ldots, u_{n-p})$ et $\mathbb{S}_p = \operatorname{vect} \mathcal{G}_p$. La matrice $A'_p \in \mathcal{M}_{n-p}(\mathbb{K})$ est en fait matrice dans la base $\mathcal{G}_p$ de l'endomorphisme de $\mathbb{S}_p$ égal à $q \circ u|_{\mathbb{S}_p}$, où q est le projecteur sur $\mathbb{S}_p$ parallèlement à $\mathbb{E}_p$. Elle possède un polynôme caractéristique scindé[16] (puisque $\chi_A(X) = \chi_{T_p}(X)\chi_{A'_p}(X)$ est scindé : on connaît le spectre de A'_p) donc au moins un vecteur propre

$$Y = \begin{pmatrix} m_1 \\ \vdots \\ m_{n-p} \end{pmatrix} \text{ associé à } \alpha \in \{\alpha_1, \ldots, \alpha_n\}$$

Le vecteur $e_{p+1} = \sum_{i=1}^{n-p} m_i u_i$ vérifie

$$u(e_{p+1}) = q \circ u(e_{p+1}) + (id_{\mathbb{E}} - q)(e_{p+1}) = \alpha e_{p+1} + v_{p+1} \quad \text{avec } v_{p+1} \in \mathbb{E}_p$$

- On pose alors $\mathcal{F}_{p+1} = (e_1, \ldots, e_{p+1})$ et $\mathbb{E}_{p+1} = \operatorname{vect} \mathcal{F}_{p+1}$. Ce sous-espace est u-stable car on vient de voir que $u(e_{p+1}) \in \mathbb{E}_{p+1}$ (la propriété (1) est donc aussi vérifiée pour $i = p+1$). On complète la famille $\mathcal{F}_{p+1}$ en une base $\mathcal{B}_{p+1} = \mathcal{F}_{p+1} \cup \mathcal{G}_{p+1}$ de $\mathbb{K}^n$ (avec, si l'on veut, $\mathcal{G}_{p+1}$ extraite de $\mathcal{G}_p$). On a alors

$$A_{p+1} = M(\Phi_A, \mathcal{B}_{p+1} = \mathcal{F}_p \cup \{e_{p+1}\} \cup \mathcal{G}_{p+1}) = \begin{pmatrix} T_p & C & D \\ 0 & \alpha & E \\ 0 & 0 & A'_{p+1} \end{pmatrix}$$

où C est une colonne $(p, 1)$, E une ligne $(1, n-p-1)$ et $A'_{p+1} \in \mathcal{M}_{n-p-1}(\mathbb{K})$ (la matrice T_p n'a pas changé, puisqu'on n'a pas changé la base du s.e.v. stable $\mathbb{E}_p$). Elle peut donc s'écrire, avec un premier bloc de taille $p+1$ triangulaire supérieur T_{p+1}

$$A_{p+1} = M(\Phi_A, \mathcal{B}_{p+1} = \mathcal{F}_{p+1} \cup \mathcal{G}_{p+1}) = \begin{pmatrix} T_{p+1} & B_{p+1} \\ 0 & A'_{p+1} \end{pmatrix}$$

On a donc remplacé p par $p+1$. Il suffit à présent de recommencer le travail avec A'_{p+1}...

[16]Si l'on voulait seulement prouver l'existence d'une forme triangulaire réduite par récurrence sur la taille de la matrice, on appliquerait l'hypothèse de récurrence au bloc A'_p, avec l'existence d'une matrice de passage

$$P_p \in \mathcal{GL}_{n-p}(\mathbb{K}) \quad \text{telle que } P_p^{-1} A'_p P_p = T''_p \text{ soit triangulaire supérieure}$$

En posant

$$P = \begin{pmatrix} I_p & 0 \\ 0 & P_p \end{pmatrix}$$

on aurait

$$P^{-1} A'_p P = \begin{pmatrix} I_p & 0 \\ 0 & P_p^{-1} \end{pmatrix} \begin{pmatrix} T_p & B_p \\ 0 & A'_p \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & P_p \end{pmatrix} = \begin{pmatrix} T_p & T_p B_p \\ 0 & T''_p \end{pmatrix}$$

ce qui démontrerait le résultat.

Prenons un exemple :

$$A = \begin{pmatrix} -2 & -1 & 2 \\ -15 & -6 & 11 \\ -14 & -6 & 11 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

vérifie (calcul simple) $\chi_A(X) = (1 - X)^3$, scindé. A est donc $\mathbb{R}$-trigonalisable et ne possède qu'une seule valeur propre 1 de multiplicité 3. La matrice $A - I_3$ est clairement de rang 2, donc le sous-espace propre est de dimension 1. La résolution du système $(A - I_3)X = 0$ donne une base de cet espace (faire les calculs ...) :

$$X_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

(avec les notations de la démonstration précédente, ici $p = 1$). Complétons avec deux vecteurs de la base canonique pour obtenir la base

$$\mathcal{B}_1 = \left(\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

La matrice de Φ_A dans cette base est

$$A_1 = \begin{pmatrix} 1 & -1 & 2 \\ 0 & -5 & 9 \\ 0 & -4 & 7 \end{pmatrix}$$

Cette matrice peut se calculer par similitude, ce qui oblige à calculer l'inverse d'une matrice de passage. On peut aussi remarquer que l'image du deuxième vecteur de base a pour coordonnées dans la base canonique $(-1, -6, -6)$. Sa première coordonnée dans $\mathcal{B}_1$ est donc forcément -1. Il suffit d'écrire

$$(-1, -6, -6) = -(1, 1, 2) + (0, -5, -4)$$

pour remplir la deuxième colonne de cette matrice. La troisième s'obtient de façon analogue. Pour être rassuré, vérifier que trace et déterminant sont corrects (on connaît $\det A$ sans calcul supplémentaire).

Toujours avec les notations précédentes, on a

$$A_1' = \begin{pmatrix} -5 & 9 \\ -4 & 7 \end{pmatrix}$$

dont le spectre est évidemment réduit à $\{1\}$ et qui possède le vecteur propre

$$Y = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

On prendra donc comme base

$$\mathcal{B}_2 = \left(\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, 3\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Dans cette base, la matrice de Φ_A s'écrit

$$T = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

(les 1 sur la diagonale ne sont pas surprenants !), matrice dont les deuxième et troisième colonnes se remplissent comme précédemment. Par exemple, l'image de $\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$ dans la base canonique s'écrit

$$\begin{pmatrix} -2 & -1 & 2 \\ -15 & -6 & 11 \\ -14 & -6 & 11 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix} = 1 \times \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Avec

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 2 & 2 & 1 \end{pmatrix}$$

on vérifie

$$P^{-1}AP = T$$

EXERCICE 5-6.6 La matrice A précédente vérifie (Cayley-Hamilton)

$$(A - I_3)^3 = 0$$

En déduire une méthode simple de calcul de A^p pour $p \in \mathbb{N}$.

5-6.3 Compléments

5-6.3.1 Endomorphisme nilpotent

DÉFINITION 5-6.7 *$u \in \mathcal{L}_{\mathbb{K}}(\mathbf{E})$ est dit* ***nilpotent*** *si et seulement s'il existe un entier $p \geqslant 1$ avec*

$$u^p = 0_{\mathcal{L}_{\mathbb{K}}(\mathbf{E})}$$

Le plus petit entier p vérifiant cette égalité est appelé indice de nilpotence de u.

EXERCICE 5-6.8 Si p est l'indice de nilpotence de u et $x \notin \ker u^{p-1}$, montrer que

$$\left(x, u(x), \ldots, u^{p-1}(x)\right) \text{ est une famille libre}$$

En déduire que, si $\mathbf{E}$ est de dimension finie n, on a forcément $p \leqslant n$. Quel raisonnement sur les polynômes annulateurs amènerait à la même conclusion ?

EXERCICE 5-6.9 Si u est nilpotent d'indice p, montrer que

$$\{0_{\mathbf{E}}\} \subsetneq \ker u \subsetneq \ker u^2 \subsetneq \cdots \subsetneq \ker u^{p-1}$$

EXERCICE 5-6.10 Si $\dim \mathbb{E} = n$ et u est nilpotent d'indice n, montrer qu'il existe une base de $\mathbb{E}$ où la matrice de u s'écrit

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Indication : choisir un x_0 avec $u^{n-1}(x_0) \neq 0_{\mathbb{E}}$ et considérer la famille $\left(u^{n-i}(x_0)\right)_{1 \leqslant i \leqslant n}$

Utiliser ce résultat pour trigonaliser facilement la matrice A étudiée à la section précédente

$$A = \begin{pmatrix} -2 & -1 & 2 \\ -15 & -6 & 11 \\ -14 & -6 & 11 \end{pmatrix}$$

EXERCICE 5-6.11 Soit $\mathbb{E}$ un $\mathbb{K}$-ev de dimension n et $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$. Montrer l'équivalence des propriétés :

i) u est nilpotent.
ii) $\chi_u(X) = (-X)^n$
iii) Il existe une base de $\mathbb{E}$ où la matrice de u est triangulaire supérieure stricte.

5-6.3.2 Spectre de $P(u)$, avec u trigonalisable

On sait déjà (*cf.* théorème **5-4**.12) que, pour $P \in \mathbb{K}[X]$ et $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E})$

$$P(\operatorname{sp} u) \subset \operatorname{sp} P(u)$$

EXERCICE 5-6.12 Montrer que, si u est un endomorphisme trigonalisable d'un $\mathbb{K}$-espace de dimension finie

$$\forall P \in \mathbb{K}[X] \qquad P(\operatorname{sp} u) = \operatorname{sp} P(u)$$

5-6.3.3 Sous-espaces caractéristiques. Décomposition diagonale par blocs triangulaires

DÉFINITION 5-6.13 *Soit $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$, et soit $\lambda \in \mathbb{K}$ une valeur propre de u, de multiplicité $m = \operatorname{mul}_u(\lambda)$. On appelle sous-espace caractéristique associé à λ le sous-espace*

$$\mathcal{C}_u(\lambda) = \ker\left(u - \lambda id_{\mathbb{E}}\right)^m$$

Propriétés :

1. **$C_u(\lambda)$ est un sous-espace u-stable, qui contient $E_u(\lambda)$.**

2. **Si on note u_λ l'endomorphisme induit sur $C_u(\lambda)$ par u, $u_\lambda - \lambda id_{C_u(\lambda)}$ est nilpotent d'indice $n_u(\lambda)$ inférieur ou égal à $\mathrm{mul}_u(\lambda)$.**

3. **La dimension de $C_u(\lambda)$ est égale à $\mathrm{mul}_u(\lambda)$**

Les deux premières propriétés sont presque évidentes et conséquences directes de la définition de $C_u(\lambda)$. Démontrons la troisième : le polynôme caractéristique de u s'écrit

$$\chi_u(X) = (X - \lambda)^m P(X)$$

avec, par définition de m, $P(\lambda) \neq 0$. Donc $\mathrm{PGCD}(P(X), (X-\lambda)^m) = 1$. Le théorème de Cayley-Hamilton et le théorème de décomposition des noyaux donnent alors

$$\mathbb{E}_n = C_u(\lambda) \oplus \ker P(u)$$

décomposition de $\mathbf{E}_n$ en deux sous-espaces u-stables. Notons u_λ et v les endomorphismes induits par u sur ces s.e.v. Le polynôme $(X-\lambda)^m$ est annulateur pour u_λ, et par application du corollaire **5-4**.13

$$\mathrm{sp}\, u_\lambda \subset \{\lambda\}$$

De même, P est annulateur pour v, et comme $P(\lambda) \neq 0$

$$\lambda \notin \mathrm{sp}\, v$$

La décomposition de $\mathbb{E}_n$ en somme directe de s.e.v. stables par u donne

$$\chi_u(X) = (X-\lambda)^m P(X) = \chi_{u_\lambda}(X)\chi_v(X)$$

Comme λ n'est pas racine de $\chi_v(X)$, on a

$(X-\lambda)^m$ divise $\chi_{u_\lambda}(X)$, ce qui implique $d^\circ\chi_{u_\lambda}(X) = \dim C_u(\lambda) \geqslant m = \underset{u}{\mathrm{mul}}(\lambda)$

De plus, comme $(X-\lambda)^m$ est annulateur pour u_λ, ce dernier est trigonalisable avec pour seule valeur propre λ. On a donc évidemment

$$\chi_{u_\lambda}(X) = (\lambda - X)^{\dim C_u(\lambda)} \text{ divise } \chi_u(X)$$

et par définition de la multiplicité

$$\dim C_u(\lambda) \leqslant \underset{u}{\mathrm{mul}}(\lambda)$$

ce qui démontre finalement la propriété 3.

REMARQUE 5-6.14 On a vu qu'une condition nécessaire pour que u soit diagonalisable était que la dimension du sous-espace propre $\mathbb{E}_u(\lambda)$ soit égale à $\operatorname{mul}_u(\lambda)$. Comme $\mathbb{E}_u(\lambda)$ est un sous-espace de $\mathcal{C}_u(\lambda)$ lui-même de dimension $\operatorname{mul}_u(\lambda)$, on a

$$\dim \mathbb{E}_u(\lambda) = \underset{u}{\operatorname{mul}}(\lambda) \Leftrightarrow \mathbb{E}_u(\lambda) = \mathcal{C}_u(\lambda) \Leftrightarrow u_\lambda = \lambda id_{\mathcal{C}_u(\lambda)}$$

Ce qui empêche d'avoir cette égalité (et interdit donc à u d'être diagonalisable), c'est qu'on a en général une écriture de u_λ de la forme

$$u_\lambda = \lambda id_{\mathcal{C}_u(\lambda)} + n_\lambda$$

avec n_λ **endomorphisme nilpotent de** $\mathcal{C}_u(\lambda)$ **non nul**, donc d'indice de nilpotence $n_u(\lambda)$ vérifiant

$$1 < n_u(\lambda) \leqslant \underset{u}{\operatorname{mul}}(\lambda)$$

Comme conséquence de l'étude précédente :

Il existe une base $\mathcal{B}_\lambda$ de $\mathcal{C}_u(\lambda)$ avec

$$M(u_\lambda, \mathcal{B}_\lambda) = \begin{pmatrix} \lambda & ? & ? & ? \\ 0 & \lambda & ? & ? \\ \vdots & 0 & \ddots & ? \\ 0 & \cdots & 0 & \lambda \end{pmatrix} = \lambda I_{\operatorname{mul}_u(\lambda)} + M_\lambda$$

avec $M_\lambda \in \mathcal{M}_{\operatorname{mul}_u(\lambda)}(\mathbb{K})$ triangulaire supérieure stricte.

Lorsque le polynôme caractéristique de $u \in \mathcal{L}_{\mathbb{K}}(\mathbb{E}_n)$ est **scindé** dans $\mathbb{K}[X]$, on pourra raisonner comme précédemment dans chaque sous-espace caractéristique, pour obtenir :

Décomposition diagonale par blocs triangulaires :

Si $\chi_u(X)$ est scindé (ce qui est toujours le cas si $\mathbb{K} = \mathbb{C}$) et si $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres distinctes de u

$$\chi_u(X) = \prod_{i=1}^{p} (\lambda_i - X)^{\operatorname{mul}_u(\lambda_i)}$$

le théorème de Cayley-Hamilton [17] donne immédiatement

$$\mathbb{E}_n = \bigoplus_{i=1}^{p} \mathcal{C}_u(\lambda_i)$$

[17] ou simplement le fait que

$$\dim \bigoplus_{i=1}^{p} \mathcal{C}_u(\lambda_i) = \sum_{i=1}^{p} \dim \mathcal{C}_u(\lambda_i) = \sum_{i=1}^{p} \underset{u}{\operatorname{mul}}(\lambda_i) = n$$

décomposition de l'espace en s.e.v. stables (de dimensions respectives $\text{mul}_u(\lambda_i)$), sur lesquels u induit des endomorphismes

$$u_{\lambda_i} = \lambda_i id_{C_u(\lambda_i)} + n_{\lambda_i}$$

avec les n_{λ_i} nilpotents. Si pour tout i on choisit comme précédemment une base $\mathcal{B}_{\lambda_i}$ qui trigonalise n_{λ_i} (ou u_{λ_i}, c'est la même chose), dans la base $\mathcal{B} = \cup \mathcal{B}_{\lambda_i}$ on aura

$$M(u,\mathcal{B}) = T = \begin{pmatrix} T_1 & 0 & 0 & 0 \\ 0 & T_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & T_p \end{pmatrix}$$

avec T_i triangulaire supérieure de taille $\text{mul}_u(\lambda_i)$ ayant une diagonale principale ne contenant que des λ_i. Ce résultat est plus précis que celui du théorème **5-6**.4. On dit qu'on a effectué une réduction de u par blocs diagonaux triangulaires supérieurs. On a évidemment un résultat analogue pour les matrices et notamment[18] : **toute matrice carrée complexe est semblable à une matrice diagonale par blocs du type précédent.**

On verra l'intérêt d'une telle réduction lors du calcul de l'exponentielle d'une matrice. Limitons nous pour le moment au calcul d'une puissance de la matrice. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et supposons que l'on ait, pour chaque valeur propre de A, déterminé une base du sous-espace caractéristique correspondant (pas nécessairement une base de trigonalisation du type $\mathcal{B}_\lambda$ évoqué plus haut). On obtient ainsi une matrice de passage P telle que

$$P^{-1}AP = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_p \end{pmatrix}$$

(écriture qui traduit la stabilité de chacun des sous-espaces caractéristiques par l'endomorphisme Φ_A). On a évidemment

$$\forall k \in \mathbb{N}^* \quad P^{-1}A^kP = \begin{pmatrix} A_1^k & 0 & 0 & 0 \\ 0 & A_2^k & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_p^k \end{pmatrix}$$

Quel est l'intérêt d'avoir travaillé avec les sous-espaces caractéristiques ? C'est que l'on sait que A_i ne possède qu'une valeur propre, disons λ_i, et que plus précisément

$$A_i = \lambda_i I_{\text{mul}_u(\lambda_i)} + N_{\lambda_i} \quad \text{avec } N_{\lambda_i} \text{ nilpotente.}$$

Le calcul de A_i^k peut alors simplement se faire avec la formule du binôme (*cf.* exercice **5-6**.6).

[18] On pourrait encore, en affinant l'étude, simplifier la forme des blocs T_i (réduite de Jordan) mais nous n'aborderons pas ce résultat ici.

EXERCICE 5-6.15 Montrer que tout endomorphisme dont le polynôme caractéristique est scindé peut se décomposer sous la forme

$$u = d + n$$

avec d diagonalisable et n nilpotent vérifiant $dn = nd$ (décomposition dite de Dunford-Jordan). On peut démontrer, mais c'est plus difficile, qu'il y a unicité d'une telle décomposition (on peut, par exemple, montrer que l'étude de cette section donne une décomposition avec d (et donc n) dans $\mathbb{K}[u]$ et déduire ainsi l'unicité).

EXERCICE 5-6.16 Une suite de matrices $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{M}_p(\mathbb{C})$ est dite convergente ssi toutes les suites de coefficients $(a_{ij}(n))_{n\in\mathbb{N}}$ sont convergentes. La limite de cette suite de matrices est alors par définition

$$\lim_{n\to+\infty} A_n = \left(\lim_{n\to\infty} a_{ij}(n)\right)_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant p}}$$

Si $A \in \mathcal{M}_p(\mathbb{C})$, montrer que la suite $(A^n)_{n\in\mathbb{N}}$ converge vers la matrice nulle ssi

$$\forall \lambda \in \operatorname{sp} A \quad |\lambda| < 1$$

EXERCICE 5-6.17 Montrer, avec les notations de cette section, que le polynôme minimal d'un endomorphisme u dont le polynôme caractéristique est scindé est

$$\pi_u(X) = \prod_{\lambda\in\operatorname{sp} u} (X-\lambda)^{n_u(\lambda)}$$

5-7 Etude des suites récurrentes linéaires à coefficients constants

Nous nous placerons dans cette section dans l'espace vectoriel $\mathbb{E} = \mathbb{C}^{\mathbb{N}}$ des suites complexes. La même étude pourrait être menée dans $\mathbb{K}^{\mathbb{N}}$ avec $\mathbb{K}$ corps quelconque de caractéristique nulle, en rajoutant une hypothèse stipulant que les polynômes intervenant dans la discussion sont scindés dans $\mathbb{K}[X]$. Nous spécifierons cependant le cas important $\mathbb{K} = \mathbb{R}$ à la fin de cette section.

5-7.1 Suites récurrentes linéaires

DÉFINITION 5-7.1 *Soit $p \in \mathbb{N}^*$. Une suite $u = (u_n)_{n\in\mathbb{N}} \in \mathbb{E}$ est dite vérifier une relation de récurrence linéaire d'ordre p à coefficients constants si et seulement s'il existe des scalaires $(\alpha_i)_{1\leqslant i\leqslant p}$ avec $\alpha_p \neq 0$ tels que*

$$\forall n \geqslant p \quad u_n = \alpha_1 u_{n-1} + \alpha_2 u_{n-2} + \cdots + \alpha_p u_{n-p} \qquad (*)$$

THÉORÈME 5-7.2 Si les scalaires $(\alpha_i)_{1\leqslant i\leqslant p}$ sont donnés, avec $\alpha_p \neq 0$, l'ensemble des suites vérifiant la relation $(*)$ est un sous-espace vectoriel de $\mathbb{E}$ de dimension p.

Démonstration : Soit $\mathcal{R}$ cet ensemble. Il est clair qu'il s'agit d'un s.e.v. De plus, si $(x_i)_{0\leqslant i\leqslant p-1}$ sont p complexes quelconques, il existe une unique suite de $\mathcal{R}$ $u = (u_n)_{n\in\mathbb{N}}$ vérifiant

$$\forall i \in \{0,\dots,p-1\} \qquad u_i = x_i$$

Il en résulte clairement que l'application

$$\mathcal{R} \to \mathbb{C}^p \text{ définie par } u = (u_n)_{n\in\mathbb{N}} \mapsto (u_0,\dots,u_{p-1})$$

est un isomorphisme d'espaces vectoriels. ■

Le but de cette section est d'obtenir explicitement, pour une suite arbitraire $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{R}$, u_n comme fonction de n. Pour ce faire, l'idée est d'obtenir une base "simple" de $\mathcal{R}$. Cela revient bien évidemment à prendre une "bonne" base de $\mathbb{C}^p$ et à "remonter" celle-ci par l'isomorphisme précédent. On pense immédiatement à la base canonique de $\mathbb{C}^p$, mais ce n'est pas une bonne idée : cette base n'est pas adaptée au problème que l'on cherche à résoudre. Le théorème de décomposition des noyaux nous permettra de répondre à la question. Commençons par quelques remarques générales.

5-7.2 Ecriture matricielle. Suites géométriques solutions

Comme la relation de récurrence (*) porte sur p termes consécutifs de la suite, une suite de $\mathcal{R}$ est caractérisée par le vecteur

$$U_0 = \begin{pmatrix} u_0 \\ \vdots \\ u_{p-1} \end{pmatrix}$$

Il est donc naturel de passer dans l'espace "des phases" $\mathbb{C}^p$ en posant

$$U_n = \begin{pmatrix} u_n \\ \vdots \\ u_{n+p-1} \end{pmatrix}$$

et la relation de récurrence (*) peut alors s'écrire

$$\forall n \in \mathbb{N} \qquad U_{n+1} = AU_n$$

où A est la matrice constante de $\mathcal{M}_p(\mathbb{C})$

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ \alpha_p & \alpha_{p-1} & \cdots & \alpha_2 & \alpha_1 \end{pmatrix}$$

et l'on en déduit que

$$\forall n \in \mathbb{N} \qquad U_n = A^n U_0$$

Cette relation donne évidemment u_n en fonction de n et des p premiers termes de la suite u. Il suffit en effet de ne retenir que le premier coefficient[19] du vecteur colonne U_n. Expliciter cette fonction revient pratiquement à exprimer A^n en fonction de n. On sait que cela peut se faire en tentant de réduire la matrice A.

EXERCICE 5-7.3 Montrer que le polynôme caractéristique[20] de la matrice A vérifie

$$(-1)^p \chi_A(X) = X^p - \alpha_1 X^{p-1} - \alpha_2 X^{p-2} - \cdots - \alpha_{p-1} X - \alpha_p$$

Quel est le polynôme minimal de A ? Montrer que A est diagonalisable ssi ce polynôme de $\mathbb{C}[X]$ a toutes ses racines simples.

Le polynôme introduit dans l'exercice précédent peut être relié d'une autre manière aux suites vérifiant la relation $(*)$, par l'intermédiaire de l'**équation caractéristique associée à la relation de récurrence** $(*)$:

THÉORÈME 5-7.4 La suite géométrique de raison $r \neq 0$ définie par

$$\forall n \in \mathbb{N} \qquad u_n = r^n$$

est dans $\mathcal{R}$ si et seulement si r est solution de l'équation

$$r^p - \alpha_1 r^{p-1} - \alpha_2 r^{p-2} - \cdots - \alpha_{p-1} r - \alpha_p = 0$$

Démonstration : évident.

COROLLAIRE 5-7.5 Lorsque l'équation (caractéristique) précédente possède p racines distinctes (non nulles puisque $\alpha_p \neq 0$) $(r_i)_{1 \leqslant i \leqslant p}$, les p suites géométriques associées

$$u^i = (r_i^n)_{n \in \mathbb{N}}$$

forment une base de $\mathcal{R}$. Si $u = (u_n)_{n \in \mathbb{N}}$ est une suite quelconque de $\mathcal{R}$, il existe des scalaires $(\mu_i)_{1 \leqslant i \leqslant p}$ tels que

$$\forall n \in \mathbb{N} \qquad u_n = \sum_{i=1}^{p} \mu_i r_i^n$$

---

[19] Plus précisément, les p suites formées par les coefficients des premières lignes des matrices de la suite $(A^n)_{n \in \mathbb{N}}$ forment une base de $\mathcal{R}$. Cette base a été évoquée plus haut, c'est celle qui correspond à la base canonique de $\mathbb{C}^p$ par l'isomorphisme "naturel" $\mathcal{R} \to \mathbb{C}^p$.

[20] A l'inverse, si $P = X^m + \lambda_1 X^{m-1} + \cdots + \lambda_{m-1} X + \lambda_m \in \mathbb{K}[X]$ est un polynôme normalisé de degré m, la matrice

$$C(P) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\lambda_m \\ 1 & 0 & 0 & \vdots & -\lambda_{m-1} \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & 0 & -\lambda_2 \\ 0 & \cdots & 0 & 1 & -\lambda_1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

est appelée matrice compagnon du polynome P. Son polynôme caractéristique est, au signe près, le polynôme P.

Démonstration : Ces p suites géométriques sont dans l'espace $\mathcal{R}$ de dimension p. Il suffit de vérifier qu'elles forment un système libre dans cet espace, ce qui se vérifie à l'aide de l'isomorphisme "naturel" $\mathcal{R} \to \mathbb{C}^p$. Les p vecteurs colonnes correspondant aux "conditions initiales" de ces p suites sont libres dans $\mathbb{C}^p$, puisque le déterminant de la matrice formée par ces colonnes est le déterminant de Vandermonde (*cf.* proposition **4-4**.41)

$$V(r_1, \dots, r_p) \neq 0 \quad \blacksquare$$

Les constantes μ_i peuvent alors être déterminées en résolvant le système de Cramer

$$\begin{cases} u_0 = \sum_{i=1}^{p} \mu_i \\ u_1 = \sum_{i=1}^{p} \mu_i r_i \\ \vdots \\ u_{p-1} = \sum_{i=1}^{p} \mu_i r_i^{p-1} \end{cases}$$

Dans le cas général (équation caractéristique n'ayant pas p racines simples) le théorème de décomposition des noyaux donne une base de $\mathcal{R}$:

5-7.3 Cas général

Considérons l'endomorphisme θ (de décalage des indices) de $\mathbb{E}$ défini par

$$u = (u_n)_{n \in \mathbb{N}} \mapsto \theta(u) = v = (u_{n+1})_{n \in \mathbb{N}}$$

Il est clair que les suites géométriques sont exactement les vecteurs propres de cet endomorphisme. Plus précisément

$$\forall r \neq 0 \qquad \ker(\theta - r id_{\mathbb{E}}) = \operatorname{vect}\left(w = (r^n)_{n \in \mathbb{N}}\right) \quad .$$

La relation de récurrence $(*)$ pouvant s'écrire aussi

$$\forall n \in \mathbb{N} \qquad u_{n+p} - (\alpha_1 u_{n+p-1} + \alpha_2 u_{n+p-2} + \dots + \alpha_p u_n) = 0$$

on a clairement

$$u \in \mathcal{R} \Leftrightarrow \theta^p(u) - \left(\alpha_1 \theta^{p-1}(u) + \alpha_2 \theta^{p-2}(u) + \dots + \alpha_{p-1}\theta(u) - \alpha_p u\right) = 0_{\mathbb{E}}$$

En d'autres termes

$$\mathcal{R} = \ker P(\theta)$$

avec $P(X) = X^p - \alpha_1 X^{p-1} - \alpha_2 X^{p-2} - \cdots - \alpha_{p-1} X - \alpha_p$. P est le polynôme intervenant dans l'équation caractéristique (et au signe près, le polynôme caractéristique de la matrice A introduite plus haut). Dans $\mathbb{C}[X]$, ce polynôme est scindé, et si $(r_i)_{1 \leqslant i \leqslant m}$ sont les racines **distinctes** de ce polynôme

$$P(X) = \prod_{i=1}^{m} (X - r_i)^{n_i}$$

En conséquence

$$\mathcal{R} = \ker \prod_{i=1}^{m} (\theta - r_i id_{\mathbb{E}})^{n_i} = \bigoplus_{i=1}^{m} \ker (\theta - r_i id_{\mathbb{E}})^{n_i}$$

par décomposition des noyaux. Lorsque $m = p$ (et donc tous les n_i valent 1) on retrouve la situation du corollaire **5-7.5**. Dans le cas général, pour trouver une base de $\mathcal{R}$, il suffit de trouver, pour chaque valeur de i, une base de $\mathcal{R}_i = \ker (\theta - r_i id_{\mathbb{E}})^{n_i}$. Ce dernier est un s.e.v. de dimension n_i (puisque c'est l'ensemble des suites complexes vérifiant une récurrence linéaire d'ordre n_i). Si $n_i = 1$, il est formé des suites géométriques de raison r_i. Sinon, considérons une suite $u \in \mathbb{E}$ dont le terme général est de la forme

$$u_n = Q(n) r_i^n \text{ avec } Q \in \mathbb{C}[X] \text{ de degré } s \geqslant 1$$

et évaluons $v = (\theta - r_i id_{\mathbb{E}})(u)$. On a

$$\forall n \in \mathbb{N} \qquad v_n = Q(n+1) r_i^{n+1} - r_i Q(n) r_i^n = (Q(n+1) - Q(n)) r_i^{n+1}$$

Il s'agit d'une suite d'un type analogue à celui de u, le polynôme Q ayant simplement été remplacé par le polynôme $Q_1(X) = r_i (Q(X+1) - Q(X))$, dont il est facile de voir qu'il est exactement de degré $s - 1$. Si donc on itère $s + 1$ fois l'endomorphisme $(\theta - r_i id_{\mathbb{E}})$ sur la suite u, le résultat sera nul.
Les n_i suites (dont on vérifie immédiatement l'indépendance)

$$(r_i^n)_{n \in \mathbb{N}}, \ (n r_i^n)_{n \in \mathbb{N}}, \ldots, (n^{n_i - 1} r_i^n)_{n \in \mathbb{N}}$$

forment donc une base de $\mathcal{R}_i = \ker (\theta - r_i id_{\mathbb{E}})^{n_i}$. La réunion de ces bases forme une base de $\mathcal{R}$. En conclusion :

PROPOSITION 5-7.6 Si $(r_i)_{1 \leqslant i \leqslant m}$ sont les racines distinctes de l'équation caractéristique associée à la récurrence linéaire $(*)$, de multiplicités respectives $(n_i)_{1 \leqslant i \leqslant m}$, les suites vérifiant cette récurrence sont exactement celles dont le terme général peut s'écrire

$$u_n = \sum_{i=1}^{m} Q_i(n) r_i^n \text{ avec les } Q_i \in \mathbb{C}_{n_i - 1}[X] \text{ quelconques.}$$

On pourra trouver ces polynômes pour une suite donnée par la méthode des coefficients indéterminés, en résolvant le système (de Cramer) associé aux conditions initiales (p équations, p inconnues).

5-7.4 Cas où $\mathbb{K} = \mathbb{R}$

On travaille à présent dans $\mathbb{E} = \mathbb{R}^{\mathbb{N}}$, avec une relation $(*)$ en supposant évidemment les $\alpha_i \in \mathbb{R}$. Notons $\mathcal{R}_{\mathbb{R}}$ l'espace des suites réelles vérifiant $(*)$ et $\mathcal{R}_{\mathbb{C}}$ le $\mathbb{C}$-s.e.v de $\mathbb{F} = \mathbb{C}^{\mathbb{N}}$ associé à la même relation de récurrence. On montre ici qu'il existe en fait une base formée de suites réelles pour l'espace $\mathcal{R}_{\mathbb{C}}$. Ces mêmes suites formeront évidemment une base de l'espace réel que l'on étudie (puisque ces p suites étant indépendantes sur $\mathbb{C}$ le seront *a fortiori* sur $\mathbb{R}$, avec $p = \dim_{\mathbb{R}} \mathcal{R}_{\mathbb{R}}$) :

- Si r_k est une racine réelle de l'équation caractéristique, de multiplicité n_k, la base correspondante de $\ker(\theta - r_k id_{\mathbb{F}})^{n_k}$

  $$(r_k^n)_{n\in\mathbb{N}},\ (nr_k^n)_{n\in\mathbb{N}},\dots,\left(n^{n_k-1}r_k^n\right)_{n\in\mathbb{N}}$$

  est formée de suites réelles.

- Si r_k est racine imaginaire (non réelle!) de la même équation, il en est de même de $\overline{r_k}$, avec la même multiplicité n_k, puisque le polynôme caractéristique est dans $\mathbb{R}[X]$. Une base de $\ker(\theta - r_k id_{\mathbb{F}})^{n_k} \oplus \ker(\theta - \overline{r_k} id_{\mathbb{F}})^{n_k}$ est

  $$(r_k^n)_{n\in\mathbb{N}},\ (\overline{r_k}^n)_{n\in\mathbb{N}},\ (nr_k^n)_{n\in\mathbb{N}},\ (n\overline{r_k}^n)_{n\in\mathbb{N}},\dots,\left(n^{n_k-1}r_k^n\right)_{n\in\mathbb{N}},\ \left(n^{n_k-1}\overline{r_k}^n\right)_{n\in\mathbb{N}}$$

  Si u est une suite de $\mathbb{F}$ et $\overline{u}$ est la suite conjuguée, la famille $(u,\overline{u})$ est engendre le même s.e.v. de $\mathbb{F}$ que $\left(\frac{1}{2}(u+\overline{u}) = \mathrm{Re}\, u, \frac{1}{2i}(u-\overline{u}) = \mathrm{Im}\, u\right)$. Les suites

  $$\begin{aligned}&(\mathrm{Re}\, r_k^n)_{n\in\mathbb{N}},\ (\mathrm{Im}\, r_k^n)_{n\in\mathbb{N}}\\ &\qquad (n\,\mathrm{Re}\, r_k^n)_{n\in\mathbb{N}},\ (n\,\mathrm{Im}\, r_k^n)_{n\in\mathbb{N}},\dots\\ &\qquad\qquad \dots,\ \left(n^{n_k-1}\,\mathrm{Re}\, r_k^n\right)_{n\in\mathbb{N}},\ \left(n^{n_k-1}\,\mathrm{Im}\, r_k^n\right)_{n\in\mathbb{N}}\end{aligned}$$

  forment donc une base de $\ker(\theta - r_k id_{\mathbb{F}})^{n_k} \oplus \ker(\theta - \overline{r_k} id_{\mathbb{F}})^{n_k}$.

  Si $r_k = \rho_k e^{i\beta_k}$ avec $\rho_k > 0$ et $\beta_k \in \mathbb{R} - \pi\mathbb{Z}$, on aura évidemment

  $$\mathrm{Re}\, r_k^n = \rho_k^n \cos n\beta_k \text{ et } \mathrm{Im}\, r_k^n = \rho_k^n \sin n\beta_k$$

On aurait d'ailleurs pu obtenir cette base par application de la

PROPOSITION 5-7.7 Les suites réelles vérifiant la relation $(*)$ (dont tous les coefficients sont supposés réels !) sont exactement les parties réelles (ou aussi imaginaires) des suites complexes vérifiant la même relation.

Démonstration : Avec des notations évidentes $\mathcal{R}_{\mathbb{R}} \subset \mathcal{R}_{\mathbb{C}}$ et, réciproquement, si $u \in \mathcal{R}_{\mathbb{C}}$, il est évident (parce que les coefficients de la relation de récurrence sont réels) que $\mathrm{Re}\, u$ (et aussi $\mathrm{Im}\, u$) sont dans $\mathcal{R}_{\mathbb{R}}$. ■

5-8 Exercices

EXERCICE 5-8.1 Soit $P \in \mathbb{Z}[X]$ de degré n et N le PGCD des nombres

$$P(0), P(1), \dots, P(n)$$

Montrer que N divise $P(x)$ pour tout $x \in \mathbb{Z}$.

EXERCICE 5-8.2 Soit A un anneau commutatif. Pour $x \in A$, on note (x) l'idéal engendré par x. Soit $(a, b) \in A^2$. Montrer que si $(a) + (b)$ est un idéal principal, il en est de même de $(a) \cap (b)$. L'anneau $\mathbb{Z}[X]$ est-il principal ?

EXERCICE 5-8.3 Soit $P \in \mathbb{Q}[X]$ et $\lambda \in \mathbb{C}$ une racine de P de multiplicité m. On suppose que $m > \dfrac{\deg P}{2}$. Montrer que $\lambda \in \mathbb{Q}$.

EXERCICE 5-8.4 Lemme de Gauss, critère d'irréductibilité d'Eisenstein : Si P est un polynôme de $\mathbb{Z}[X]$, on note $c(P)$ le PGCD de ses coefficients.

1. Si P_1 et P_2 sont dans $\mathbb{Z}[X]$ et si p est un diviseur premier de $c(P_1P_2)$, montrer que p divise $c(P_1)c(P_2)$. En déduire que $c(P_1P_2) = c(P_1)c(P_2)$.

2. Un polynôme P non constant de $\mathbb{Z}[X]$ est dit irréductible s'il ne peut s'écrire comme produit de polynômes de $\mathbb{Z}[X]$ dont les degrés sont tous strictement inférieurs au degré de P. Montrer qu'un polynôme de $\mathbb{Z}[X]$ est irréductible ssi il l'est dans $\mathbb{Q}[X]$.

3. Soit $P = \sum_0^n a_k X^k$ un polynôme non constant de $\mathbb{Z}[X]$. On suppose qu'il existe un nombre premier p divisant tous les a_k sauf a_n et tel que p^2 ne divise pas a_0. Montrer que P est irréductible dans $\mathbb{Q}[X]$. Application : Pour p premier, $1 + X + \dots + X^{p-1}$ est irréductible (poser $X = 1 + Y$).

EXERCICE 5-8.5 Trouver P et $Q \in \mathbb{R}[X]$ tels que $P^2 = 1 + (X^2 - 1)Q^2$.

EXERCICE 5-8.6 Soit $P = a_0 + a_1X + \dots + a_nX^n \in \mathbb{Z}[X]$ avec $a_0a_n \neq 0$. Montrer que si $r = \dfrac{p}{q}$ est une racine rationnelle de P avec $p \wedge q = 1$ alors $p|a_0$, $q|a_n$, $p - q|P(1)$ et $p + q|P(-1)$. Factoriser sur $\mathbb{Q}$ le polynôme $4X^4 - 28X^3 + 45X^2 - 6X - 18$.

EXERCICE 5-8.7 Trouver $P \in \mathbb{C}[X]$ avec $P(X^2) = P(X)P(X + 1)$.

EXERCICE 5-8.8 Soient p et q deux entiers naturels non nuls. Trouver $P \in \mathbb{C}[X]$ vérifiant $(P')^p$ divise P^q.

EXERCICE 5-8.9 Soit $P \in \mathbb{R}[X]$ scindé. Montrer que, pour $a \in \mathbb{R}$ $P + aP'$ est scindé. Soit $P \in \mathbb{C}[X]$. Montrer que toute racine de P' appartient à l'enveloppe convexe des racines de P (c'est à dire à la plus petite partie convexe de $\mathbb{C}$ contenant ces racines. On pourra utiliser la décomposition en éléments simples de $\dfrac{P'}{P}$).

EXERCICE 5-8.10 On définit la suite de Fibonacci $F_0 = 0, F_1 = 1$ et

$$F_{n+2} = F_n + F_{n+1}$$

- Si on met en oeuvre l'algorithme d'Euclide, quel est le nombre de divisions à effectuer pour obtenir le PGCD de F_{n+2} et F_{n+1} ?

- Soient a et $b \in \mathbb{N}^*$ avec $a < b$. On pose $r_0 = b$, $r_1 = a$ et on effectue l'algorithme d'Euclide : $r_0 = q_1 r_1 + r_2, \cdots, r_{n-1} = r_n q_n + r_{n+1}$, r_{n+1} étant le dernier reste non nul. On veut établir que le nombre de divisions euclidiennes est inférieur ou égal à 5 fois le nombre de chiffres de b en écriture décimale.
  Montrer que $F_{n+2} \leqslant b$ et $F_{n+1} \leqslant a$. Si b a p chiffres en écriture décimale, montrer que
  $$\left(\frac{1+\sqrt{5}}{2}\right)^n \leqslant \sqrt{5}10^p$$
  Conclure.

EXERCICE 5-8.11 On considère $\mathcal{A} = \{a + ib\sqrt{2} \mid (a,b) \in \mathbb{Z}^2\}$.

- Montrer que $\mathcal{A}$ est un sous anneau de $\mathbb{C}$.
- Montrer que, en posant $N(a + ib\sqrt{2}) = a^2 + 2b^2$, on a, pour A et $B \in \mathcal{A}$ avec $B \neq 0$ l'existence de $Q, R \in \mathcal{A}$ tels que $A = BQ + R$ avec $N(R) < N(B)$.
- On introduit la notion de nombre premier (élément irréductible) dans $\mathcal{A}$. Etablir un théorème de décomposition en facteurs premiers dans l'anneau $\mathcal{A}$.

EXERCICE 5-8.12 Soit $P \in \mathbb{Z}[X]$ avec $P(0)$ impair. Montrer qu'aucune racine de P n'est de la forme 2^r avec $r \in \mathbb{Q}^{*+}$.

EXERCICE 5-8.13 Soit $P(X) = \sum_{k=0}^{n} a_k X^k$ un polynôme de $\mathbb{R}[X]$ scindé à racines simples. Etudier le signe de $PP'' - (P')^2$ et en déduire que $\forall k \quad a_k^2 > a_{k-1}a_{k+1}$.

EXERCICE 5-8.14 Soit $\varphi : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$ non constante telle que
$$\forall X, Y \quad \varphi(X)\varphi(Y) = \varphi(XY)$$

1. Montrer que X inversible $\Leftrightarrow \varphi(X) \neq 0$.
2. On suppose de plus que $\varphi(X)$ s'exprime comme polynôme homogène de degré 2 des coefficients de X. Montrer que $n = 2$ et $\varphi(X) = \det X$.

EXERCICE 5-8.15 Soit A une matrice carrée d'ordre n complexe vérifiant
$$3A^3 = A^2 + A + I_n$$
Calculer $\lim\limits_{p \to +\infty} A^p$.

EXERCICE 5-8.16 Soit K un corps de caractéristique différente de 2 et G un sous-groupe de $GL_n(K)$ tel que, pour tout $A \in G$ on ait $A^2 = I_n$. Que peut on dire du cardinal de G ?
Etudier l'existence d'isomorphismes de groupes entre $GL_n(K)$ et $GL_m(K)$, entre $GL_n(\mathbb{R})$ et $GL_m(\mathbb{Q})$, entre $GL_n(\mathbb{R})$ et $GL_m(\mathbb{C})$.

EXERCICE 5-8.17 On considère K^n ramené à sa base canonique $\{e_1, \ldots, e_n\}$ et, pour $\sigma \in S_n$ l'endomorphisme défini par $f_\sigma(e_i) = e_{\sigma(i)}$. Déterminer $\det f_\sigma$, son polynôme caractéristique et son polynôme minimal.

EXERCICE 5-8.18 Soit E le $\mathbb{R}$-ev des applications continues de $[0,+\infty[$ dans $\mathbb{R}$ ayant une limite finie en $+\infty$. Soit $\varphi \in \mathcal{L}(E)$ qui à f associe g définie par $g(x) = f(x+1)$. Déterminer les éléments propres de φ.
Même question avec $E = C^0([0,1],\mathbb{R})$, $n \in \mathbb{N}$ et $g(x) = \dfrac{1}{x^{n+1}}\displaystyle\int_0^x t^n f(t)dt$.

EXERCICE 5-8.19 Trouver les valeurs propres des matrices de $\mathcal{M}_n(\mathbb{C})$:

$$\begin{pmatrix} & & & & 1 \\ & & & & 2 \\ & & 0 & & 3 \\ & & & & \vdots \\ 1 & 2 & 3 & \cdots & n \end{pmatrix} \begin{pmatrix} 0 & y & \cdots & y \\ x & 0 & y & \vdots \\ \vdots & x & 0 & y \\ x & \cdots & x & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ -1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & -1 & 0 & -1 \\ 0 & \cdots & 0 & -1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} x_1^2 & \cdots & x_1x_j & \cdots & x_1x_n \\ \vdots & & \vdots & & \vdots \\ x_ix_1 & \cdots & x_ix_j & \cdots & x_ix_n \\ \vdots & & \vdots & & \vdots \\ x_nx_1 & \cdots & x_nx_j & \cdots & x_n^2 \end{pmatrix}$$

Ces matrices sont-elles diagonalisables ?

EXERCICE 5-8.20 Condition nécessaire et suffisante pour que

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & \alpha_n \\ 0 & \cdots & \cdots & \cdots & \alpha_{n-1} & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & \alpha_2 & 0 & \cdots & \cdots & 0 \\ \alpha_1 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

appartenant à $\mathcal{M}_n(\mathbb{C})$ soit diagonalisable ? Même question sur $\mathbb{R}$.

EXERCICE 5-8.21 Trouver les matrices $A \in \mathcal{M}_3(\mathbb{R})$ telles que $A^2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ -8 & 2 & 4 \end{pmatrix}$

EXERCICE 5-8.22 Déterminer le commutant dans $\mathcal{M}_3(\mathbb{C})$ et les sous-espaces stables de la matrice $\begin{pmatrix} 1 & j & j^2 \\ j & j^2 & 1 \\ j^2 & 1 & j \end{pmatrix}$ $\left(j = e^{\frac{2i\pi}{3}}\right)$

EXERCICE 5-8.23 Soit A et $B \in \mathcal{M}_n(\mathbb{C})$. On suppose qu'il existe $C \in \mathcal{M}_n(\mathbb{C})$ non nulle telle que $AC = CB$. Montrer que $\forall P \in \mathbb{C}[X]$ $P(A)C = CP(B)$. En déduire que A et B ont une valeur propre commune. Etudier la réciproque.

EXERCICE 5-8.24 Soit $f \in L_K(E)$ diagonalisable. Montrer que tout sev de E stable par f possède un supplémentaire stable. Etudier la réciproque pour $K = \mathbb{C}$ puis $K = \mathbb{R}$.

EXERCICE 5-8.25 Soit E un K-ev de dimension n. Montrer que si $u \in \mathcal{L}(E)$ est diagonalisable, il existe $p \in \mathbb{N}^*$, p scalaires distincts $\alpha_1, \ldots, \alpha_p$ et p endomorphismes de E, $v_1, \ldots, v_p$ tels que

$$\forall P \in K[X] \quad P(u) = \sum_{i=1}^{p} P(\alpha_i) v_i$$

Etudier la réciproque.

EXERCICE 5-8.26 Si E est un K-ev de dimension n et $u \in \mathcal{L}(E)$, on dit que u est cyclique s'il existe $x_0 \in E$ tel que E soit engendré par $\{u^k(x_0) \mid k \in \mathbb{N}\}$. Montrer que le commutant de u dans $\mathcal{L}(E)$ est alors égal à $K[u]$. Application : trouver toutes les matrices de $\mathcal{M}_n(\mathbb{R})$ vérifiant :

$$X^2 = \begin{pmatrix} 1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

Montrer qu'un endomorphisme est cyclique ssi son polynôme minimal est égal à son polynôme caractéristique.

EXERCICE 5-8.27 Soit $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que A est nilpotente ssi, pour tout $p \in \{1, \ldots, n\}$ trace(A^p)=0. Montrer qu'alors, pour tout $M \in \mathcal{M}_n(\mathbb{C})$

$$AM = MA \Rightarrow \det(A + M) = \det M$$

Etudier la réciproque.
Soient A et $B \in \mathcal{M}_n(\mathbb{C})$ et $C = AB - BA$. On suppose que C commute avec A et B. Montrer que C est nilpotente.

EXERCICE 5-8.28 Soient A et $B \in \mathcal{M}_n(\mathbb{C})$ et f_1, f_2 les endomorphismes de $\mathcal{L}(\mathcal{M}_n(\mathbb{C}))$ définies par $f_1(X) = AX$ et $f_2(X) = XB$. Déterminer les spectres de f_1 et f_2. Comment obtient les espaces propres correspondants ? Montrer que f_1 est diagonalisable ssi A l'est.
On note $\varphi = f_1 - f_2$. Montrer que si λ_1 est valeur propre de A et λ_2 valeur propre de B alors $\lambda_1 - \lambda_2$ est valeur propre de φ. Etudier la réciproque.

EXERCICE 5-8.29 Calculer A^n avec

$$A = \begin{pmatrix} 3 & -5 & 2 & -6 \\ 0 & 5 & 0 & 4 \\ -2 & 7 & -1 & 11 \\ 0 & -4 & 0 & -3 \end{pmatrix}$$

EXERCICE 5-8.30 Soient $A_1, \cdots, A_n$ n matrices nilpotentes de $\mathcal{M}_n(K)$ commutant deux à deux. Montrer que $A_1 A_2 \cdots A_n = 0$.

EXERCICE 5-8.31 $\mathcal{M}_p(\mathbb{C})$ est muni d'une norme quelconque. Si $A \in \mathcal{M}_p(\mathbb{C})$, montrer que :

$$\lim_{m \to +\infty} A^m = 0 \iff \forall \lambda \in \operatorname{sp}(A) \ \ |\lambda| < 1$$

Pour $A \in \mathcal{M}_p(\mathbb{C})$ on note $\rho(A) = \max\{|\lambda|, \ \lambda \text{ valeur propre de } A\}$. Si $A = (a_{i,j})$ et $B = (b_{i,j}) \in \mathcal{M}_p(\mathbb{C})$ vérifient $\forall i, j \ \ a_{i,j} > 0$ et $|b_{i,j}| \leqslant a_{i,j}$, montrer que $\rho(B) \leqslant \rho(A)$.

EXERCICE 5-8.32 Soit E un espace de dimension n, et soit p un projecteur de rang r. On définit φ sur $\mathcal{L}(E)$ par : $\varphi(f) = fp - pf$. Diagonaliser φ.

EXERCICE 5-8.33 Soient $A \in \mathcal{M}_{3,2}(\mathbb{R})$ et $B \in \mathcal{M}_{2,3}(\mathbb{R})$ telles que

$$AB = \begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}$$

Calculer le rang de AB. AB est-elle diagonalisable ? Calculer le rang de BA. Calculer BA.

Chapitre 6

Espaces normés : suites et topologie

6-1 Suites réelles et complexes

6-1.1 Borne supérieure

On sait que, si $A \subset \mathbb{R}$ est une partie non vide et majorée, l'ensemble de ses majorants possède un plus petit élément, qu'on appelle **borne supérieure de** A et qu'on note $\sup(A)$. Il est à noter que $M = \sup(A)$ n'est pas en général un élément de A et est caractérisée par les deux propriétés[1] :

$$\forall x \in A \quad x \leqslant M$$

$$\forall \epsilon > 0 \quad \exists x \in A \quad M - \epsilon < x \leqslant M$$

La première propriété traduit le fait que M majore A, la seconde qu'un réel strictement inférieur à M n'est pas majorant de A.

Le fait d'admettre que $(\mathbb{R}, +, \times, \leqslant)$ est un corps totalement ordonné possédant cette propriété de la borne supérieure a de nombreuses conséquences. Citons notamment :

[1] Cette caractérisation de la borne supérieure utilise le fait que l'ordre est total sur $\mathbb{R}$. Dans un ensemble partiellement ordonné $(X, \leqslant)$, une partie non vide A possède une borne supérieure M si et seulement si M majore A et si tout majorant M' de A vérifie $M \leqslant M'$.

- $\mathbb{N}$ est non majoré dans $\mathbb{R}$, ($\mathbb{R}$ est "archimédien"), ce qui peut se traduire par[2]

$$\forall x \in \mathbb{R}^{*+} \quad \forall y \in \mathbb{R} \quad \exists n \in \mathbb{N} \quad nx > y$$

  L'existence d'une partie entière pour tout réel en découle.

- Si b est un entier supérieur ou égal à 2, tout réel admet un développement illimité en base b (décimal si $b = 10$, binaire si $b = 2$), unique si l'on demande à ce développement d'être *propre*, c'est-à-dire de ne pas se terminer par une suite illimitée de symboles représentant l'entier $b - 1$.

- L'existence d'une racine $p^{\text{ième}}$ pour tout réel positif, si $p \geqslant 2$ est un entier naturel (par exemple $\sqrt{2}$ peut être obtenu comme la borne supérieure de $\{x \in \mathbb{Q}^+ \mid x^2 < 2\}$).

Le théorème de convergence d'une suite monotone découle également de cette propriété. Nous nous intéresserons également à une autre conséquence de la propriété de la borne supérieure, qui donnera un critère commode de convergence des suites réelles : toute suite de Cauchy de nombres réels converge.

6-1.2 Convergence d'une suite réelle

DÉFINITION 6-1.1 (Suite convergente) *Une suite $(u_n)_{n\in\mathbb{N}}$ de réels est convergente si et seulement s'il existe un réel l vérifiant*

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \geqslant N_\varepsilon \quad |u_n - l| \leqslant \varepsilon$$

Le réel l est alors unique et on le note $l = \lim\limits_{n\to+\infty} u_n$.

THÉORÈME 6-1.2 Toute suite convergente est bornée. Le produit d'une suite bornée par une suite qui converge vers 0 est convergente vers 0.

De la définition et de ce théorème découlent les résultats sur les relations entre limites et opérations sur les suites, notamment :

THÉORÈME 6-1.3 Si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont deux suites réelles convergentes et $\alpha \in \mathbb{R}$

- **La suite $(u_n + \alpha v_n)_{n\in\mathbb{N}}$ converge et $\lim\limits_{n\to+\infty} u_n + \alpha v_n = \lim\limits_{n\to+\infty} u_n + \alpha \lim\limits_{n\to+\infty} v_n$.**

- **La suite $(u_n v_n)_{n\in\mathbb{N}}$ converge et $\lim\limits_{n\to+\infty} u_n v_n = (\lim\limits_{n\to+\infty} u_n)(\lim\limits_{n\to+\infty} v_n)$**

- **Si la suite $(u_n)_{n\in\mathbb{N}}$ a une limite non nulle, la suite de terme général $w_n = \dfrac{1}{u_n}$ est définie à partir d'un certain rang et converge vers $\dfrac{1}{\lim\limits_{n\to+\infty} u_n}$**

[2]Cela peut s'écrire aussi, avec la notion de suite convergente, $\lim \frac{1}{n} = 0$.

Démonstration : La démonstration de la première propriété est très simple ("couper les ε en 2"). Pour la seconde, si on note l et l' les limites respectives des deux suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$, il suffit d'écrire que $u_nv_n - ll' = (u_n - l)v_n + l(v_n - l')$ et de remarquer que le terme de droite se décompose à l'aide de produits de suites bornées et de suites tendant vers 0. Enfin, dire que la suite $(u_n)_{n\in\mathbb{N}}$ tend vers $l \neq 0$ entraîne l'existence d'un rang N_0 à partir duquel $|u_n| \geqslant \frac{|l|}{2}$, ce qui montre que (w_n) est définie à partir de N_0 et est bornée. On conclut en remarquant que

$$w_n - \frac{1}{l} = w_n\left(1 - \frac{u_n}{l}\right)$$

est produit d'une suite bornée par une suite tendant vers 0. ■

THÉORÈME 6-1.4 Si $(u_n)_{n\in\mathbb{N}}$ est une suite réelle croissante, elle converge si et seulement si elle est majorée. On a alors $\lim_{n\to+\infty} u_n = \sup\{u_n \mid n \in \mathbb{N}\}$.

Démonstration : Si la suite (u_n) est majorée et si on note

$$M = \sup\{u_n \mid n \in \mathbb{N}\}$$

on a, si $\varepsilon > 0$ est quelconque,

$$\exists\, n_0 \in \mathbb{N} \quad M - \varepsilon < u_{n_0} \leqslant M$$

Comme la suite est croissante, pour $n > n_0$ on a également

$$M - \varepsilon < u_{n_0} \leqslant u_n \leqslant M$$

ce qui entraîne bien la convergence de la suite vers M. Si la suite n'est pas majorée, on a

$$\forall\, A \in \mathbb{R} \quad \exists\, n_0 \in \mathbb{N} \quad u_{n_0} \geqslant A$$

On a alors *a fortiori*, pour $n \geqslant n_0$, $u_n \geqslant A$, ce qu'on traduit par la divergence de la suite vers $+\infty$. ■

EXERCICE 6-1.5 Montrer que la borne supérieure d'un sous-ensemble majoré A de $\mathbb{R}$ est le seul majorant de cette partie A qui soit limite d'une suite de points de A.

6-1.3 Valeur d'adhérence d'une suite réelle

6-1.3.1 Suite extraite

Si $u = (u_n)_{n\in\mathbb{N}}$ est une suite réelle, on appelle suite extraite de la suite u (on parle aussi de sous-suite de u) toute suite $v = (u_{n_p})_{p\in\mathbb{N}}$ obtenue en "choisissant" une suite d'entiers strictement croissante

$$n_0 < n_1 < \cdots < n_p < n_{p+1} < \cdots$$

et en ne "retenant" des termes de la suite u que les termes d'indices ainsi sélectionnés. Une définition[3] plus formelle est :

[3]Par abus de langage, on dira que $v_n = u_{n^2-12n+32}$ est une suite extraite de la suite u, bien que l'indice ne soit croissant qu'à partir du rang 6. De même $w_n = u_{n^2-10}$ n'est défini qu'à partir du rang 4.

DÉFINITION 6-1.6 (Suite extraite) *La suite $v = (v_n)_{n\in\mathbb{N}}$ est extraite de la suite $u = (u_n)_{n\in\mathbb{N}}$ si et seulement s'il existe une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que, pour tout entier n, on ait $v_n = u_{\varphi(n)}$.*

PROPOSITION 6-1.7 Si $u = (u_n)_{n\in\mathbb{N}}$ est une suite réelle convergente vers l, toute suite extraite de la suite u converge vers la même limite.

Démonstration : Il suffit de remarquer que, si $v_n = u_{\varphi(n)}$ avec $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante, on a $\varphi(n) \geqslant n$ pour tout n, et d'appliquer la définition de la convergence vers l de la suite u pour conclure immédiatement. ■

6-1.3.2 Valeur d'adhérence d'une suite réelle

DÉFINITION 6-1.8 (Valeur d'adhérence) *Si $u = (u_n)_{n\in\mathbb{N}}$ est une suite réelle et si $l \in \mathbb{R}$, on dit que l est une valeur d'adhérence de la suite u si et seulement s'il existe une suite extraite de u qui converge vers l.*

La proposition suivante est une caractérisation commode des valeurs d'adhérence :

PROPOSITION 6-1.9 Un réel l est valeur d'adhérence de la suite réelle u ssi, pour tout $\varepsilon > 0$, l'ensemble

$$\{n \in \mathbb{N} \mid u_n \in]l-\varepsilon, l+\varepsilon[\}$$

est infini.

Démonstration : On remarquera que c'est l'ensemble des indices qui est infini, l'ensemble des valeurs de la suite appartenant à l'intervalle $]l-\varepsilon, l+\varepsilon[$ pouvant être fini (penser par exemple aux suites constantes) ! Si l est limite d'une suite v extraite de la suite u, et si $\varepsilon > 0$ est donné, l'intervalle $]l-\varepsilon, l+\varepsilon[$ contient tous les termes de la suite v à partir d'un certain rang, et l'ensemble d'indices considéré est bien infini. Réciproquement, si cet ensemble est infini pour tout $\varepsilon > 0$, on construit de proche en proche les termes d'une suite extraite de la suite u qui converge vers l par le procédé suivant : On prend un indice n_0 tel que $u_{n_0} \in]l-1, l+1[$, puis $n_1 > n_0$ avec $u_{n_1} \in \left]l-\frac{1}{2}, l+\frac{1}{2}\right[$. Si on suppose construits des entiers $n_0 < n_1 < \cdots < n_p$ avec, pour tout $i \leqslant p$, $u_{n_i} \in \left]l-\frac{1}{i+1}, l+\frac{1}{i+1}\right[$, comme l'ensemble $\left\{n \in \mathbb{N} \mid u_n \in \left]l-\frac{1}{p+2}, l+\frac{1}{p+2}\right[\right\}$ est infini on pourra choisir un indice $n_{p+1} > n_p$ avec $|u_{n_{p+1}} - l| < \frac{1}{p+2}$. En posant, pour tout $p \in \mathbb{N}$, $v_p = u_{n_p}$, on obtient une suite extraite qui répond à la question. ■

La proposition **6-1**.7 montre qu'une suite convergente possède une unique valeur d'adhérence qui est égale à sa limite. La suite $u_n = (1 + (-1)^n)n$ donne un exemple de suite réelle ayant une unique valeur d'adhérence (le prouver) et qui ne converge pas.

6-1.3.3 Théorème de Bolzano-Weierstrass

THÉORÈME 6-1.10 (Bolzano-Weierstrass) De toute suite bornée de réels on peut extraire une sous-suite convergente. De manière équivalente, on peut dire que toute suite bornée de réels possède au moins une valeur d'adhérence.

Démonstration : Comme la suite est bornée, il existe deux réels m et M tels que, pour tout n on ait $m \leqslant u_n \leqslant M$. La propriété de la borne supérieure permet de définir une suite v par

$$\forall n \in \mathbb{N} \quad v_n = \sup \{u_k \mid k \geqslant n\}$$

La suite v est clairement décroissante, et vérifie

$$\forall n \in \mathbb{N} \quad m \leqslant u_n \leqslant v_n$$

Etant minorée par m, elle converge vers un réel L. On montre ensuite que L est une valeur d'adhérence de la suite u. En effet, si $\varepsilon > 0$ est un réel, on peut trouver un rang N_ε à partir duquel $L \leqslant v_n \leqslant L + \varepsilon$ (convergence de la suite v). Si $n \geqslant N_\varepsilon$ on a aussi $L - \varepsilon < v_n$, ce qui, compte tenu de la définition de v_n, entraîne l'existence d'un entier $k \geqslant n$ tel que $L - \varepsilon \leqslant u_k \leqslant v_n \leqslant L + \varepsilon$. On en déduit ainsi l'existence d'une infinité de termes de la suite u dans $]L - \varepsilon, L + \varepsilon[$ (puisque l'indice k est "aussi grand qu'on le veut"). ■

REMARQUE 6-1.11 On vient de prouver qu'une suite bornée possède toujours comme valeur d'adhérence le réel

$$\overline{\lim_{n \to +\infty}} \ u_n = \lim_{n \to +\infty} \sup_{p \geqslant n} u_p$$

qu'on appelle **limite supérieure** de la suite u. La démonstration précédente montre que si une suite extraite de la suite u est définie par $w_n = u_{\varphi(n)}$, comme on a $\varphi(n) \geqslant n$ (récurrence sur n)

$$\forall n \in \mathbb{N} \quad w_n = u_{\varphi(n)} \leqslant v_n$$

Ceci montre que, si la suite w est convergente, sa limite est forcément inférieure à $\overline{\lim\limits_{n \to +\infty}} \ u_n$. L'ensemble des valeurs d'adhérence de la suite bornée u a donc un plus grand élément qui est la limite supérieure de la suite. On définirait de même la **limite inférieure** de la suite bornée u par

$$\underline{\lim_{n \to +\infty}} \ u_n = \lim_{n \to +\infty} \inf_{p \geqslant n} u_p$$

(limite d'une suite croissante majorée), et on prouverait aisément qu'il s'agit de la plus petite valeur d'adhérence de la suite u.

EXERCICE 6-1.12 Si u et v sont deux suites réelles bornées, montrer que

$$\overline{\lim_{n \to +\infty}} \ (u_n + v_n) \leqslant \overline{\lim_{n \to +\infty}} \ u_n + \overline{\lim_{n \to +\infty}} \ v_n$$

et donner un exemple où l'inégalité est stricte. Donner de même un résultat concernant les limites inférieures. Que peut-on dire de $\overline{\lim\limits_{n \to +\infty}} \ (-u_n)$?

On a vu précédemment qu'une suite réelle convergente est bornée et ne possède qu'une seule valeur d'adhérence. Le théorème suivant, réciproque de cette propriété, est souvent utilisé pour prouver la convergence d'une suite :

THÉORÈME 6-1.13 Si une suite réelle est bornée et ne possède qu'une valeur d'adhérence l, elle est convergente vers l.

Démonstration : Si la suite u_n ne converge pas vers l on a :

$$\exists\, \varepsilon > 0 \quad \forall\, N \in \mathbb{N} \quad \exists\, n \geqslant N \quad |u_n - l| > \varepsilon$$

On peut donc extraire une suite $v_n = u_{\varphi(n)}$ vérifiant $|v_n - l| > \varepsilon$ pour tout n. La suite v est bornée et possède donc, d'après le théorème de Bolzano-Weierstrass, une valeur d'adhérence l' (qui est aussi valeur d'adhérence de la suite u, puisqu'une sous-suite d'une suite extraite de la suite u est aussi sous-suite de u) qui vérifiera $|l' - l| \geqslant \varepsilon$. Ceci contredit l'unicité de la valeur d'adhérence de u. ■

6-1.4 Suites de Cauchy

DÉFINITION 6-1.14 (Suite de Cauchy)

Cauchy!Suite Une suite réelle $u = (u_n)_{n\in\mathbb{N}}$ est dite de Cauchy si et seulement si elle vérifie la propriété :

$$\forall\, \varepsilon > 0 \quad \exists\, N_\varepsilon \in \mathbb{N} \quad \forall\, n, m \in \mathbb{N} \quad n \geqslant N_\varepsilon \textit{ et } m \geqslant N_\varepsilon \Longrightarrow |u_n - u_m| \leqslant \varepsilon$$

REMARQUE 6-1.15 Dans cette définition, les entiers n et m sont quelconques, "indépendants" si l'on veut. On peut se convaincre qu'une suite u de Cauchy vérifie $\lim\limits_{n\to+\infty} u_{n+p} - u_n = 0$, si p est un entier naturel quelconque. Cette propriété ne caractérise cependant pas les suites de Cauchy, comme l'exemple $u_n = \ln n$ le montre.

THÉORÈME 6-1.16 Toute suite convergente de réels est une suite de Cauchy.

Démonstration : Si la suite u converge vers l et si $\varepsilon > 0$ est donné, il existe un rang N à partir duquel $|u_k - l| \leqslant \frac{\varepsilon}{2}$. Si n et m sont supérieurs à N, on a immédiatement par inégalité triangulaire $|u_n - u_m| \leqslant \varepsilon$. ■

L'intérêt de la notion de suite de Cauchy réside dans la réciproque de ce théorème, qui donne un moyen de prouver qu'une suite réelle est convergente, sans déterminer la valeur de la limite *a priori*.

THÉORÈME 6-1.17 Toute suite de Cauchy de nombres réels est convergente. On dit que $(\mathbb{R}, |\,|)$ est un espace complet.

Démonstration : Une suite réelle u de Cauchy est bornée. En effet, à partir d'un rang N_1, on a une majoration $|u_n - u_{N_1}| \leqslant 1$, ce qui entraîne pour tout entier naturel n la majoration

$$|u_n| \leqslant \max(|u_0|, |u_1|, \ldots, |u_{N_1-1}|, |u_{N_1}| + 1)$$

La suite u possède donc une valeur d'adhérence, d'après le théorème de Bolzano-Weierstrass. Si on note l cette valeur d'adhérence et si ε est un réel > 0 arbitraire, il y a une infinité d'indices k vérifiant $u_k \in [l - \varepsilon, l + \varepsilon]$. De plus, on peut trouver un rang N_ε tel que

$$n \geqslant N_\varepsilon \text{ et } m \geqslant N_\varepsilon \Longrightarrow |u_n - u_m| \leqslant \varepsilon$$

Si $n \geqslant N_\varepsilon$ est un entier arbitraire, en choisissant un entier $k \geqslant N_\varepsilon$ avec $u_k \in [l - \varepsilon, l + \varepsilon]$, on aura

$$|u_n - l| \leqslant |u_n - u_k| + |u_k - l| \leqslant 2\varepsilon$$

ce qui montre la convergence de la suite u vers l. ■

6-1.5 Extension aux suites complexes

C'est la fonction "module" qui prolonge la fonction "valeur absolue" à $\mathbb{C}$, et qui permet de définir la distance de deux nombres complexes z et z' par $|z - z'|$. On peut alors étendre les définitions :

DÉFINITION 6-1.18 *Une suite $(z_n)_{n\in\mathbb{N}}$ de complexes est convergente si et seulement s'il existe un complexe l vérifiant*

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \geqslant N_\varepsilon \quad |z_n - l| \leqslant \varepsilon$$

Si $z_n = u_n + iv_n$ où u_n et v_n sont deux suites réelles, et si $l = a + ib$ avec $a, b \in \mathbf{R}$, les majorations

$$|u_n - a| \leqslant |z_n - l|, |v_n - b| \leqslant |z_n - l| \text{ et } |z_n - l| \leqslant |u_n - a| + |v_n - b|$$

montrent que la convergence de la suite z_n vers l équivaut à la convergence simultanée des suites des parties réelles et des parties imaginaires vers les parties réelle et imaginaire de l. Cependant, les calculs sur les complexes ne se ramenant pas systématiquement à des séparations en parties réelles et imaginaires, d'autres techniques (notamment l'utilisation des modules et arguments) peuvent être utilisées.

EXERCICE 6-1.19 Etudier la convergence de la suite

$$z_n = \left(1 + \frac{in}{n^2 - 1}\right)^n$$

Si la notion de suite monotone n'a évidemment aucun sens dans $\mathbb{C}$, les autres notions introduites précédemment se généralisent sans difficulté. En particulier, un complexe l est valeur d'adhérence d'une suite complexe z_n ssi pour tout $\varepsilon > 0$, $\{n \in \mathbb{N} \mid u_n \in D(l, \varepsilon[\}$ est infini, où $D(l, \varepsilon[= \{z \in \mathbb{C} \mid |z - l| < \varepsilon\}$ est le disque ouvert de centre l et de rayon ε. On a toujours le théorème de Bolzano-Weierstrass :

THÉORÈME 6-1.20 (Bolzano-Weierstrass) Toute suite bornée de complexes possède au moins une valeur d'adhérence.

Démonstration : Dire qu'une suite $z = (z_n)_{n\in\mathbb{N}}$ de complexes est bornée revient à affirmer l'existence d'un réel M tel que pour tout n on ait $|z_n| \leqslant M$. La suite $u_n = \operatorname{Re} z_n$ est donc bornée, et on peut en extraire une sous-suite $u_{\varphi(n)}$ convergente vers un réel a. La suite $v_n = \operatorname{Im} z_{\varphi(n)}$ est aussi bornée, et on peut en extraire une sous-suite $v_{\psi(n)}$ qui converge vers un réel b. La suite $z_{\varphi\circ\psi(n)}$ est extraite de la suite z et converge vers le complexe $a + ib$. ■

De même, la notion de suite de Cauchy se généralise, et les inégalités entre module et valeurs absolues des parties réelles et imaginaires rappelées au début de ce paragraphe montrent qu'une suite complexe est de Cauchy ssi les suites des parties réelles et imaginaires le sont. On en déduit le

THÉORÈME 6-1.21 Toute suite de Cauchy de complexes est convergente.

6-2 Norme sur un espace vectoriel. Distance associée

Dans toute cette partie, E est un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

6-2.1 Norme et distance

DÉFINITION 6-2.1 (Norme) *Si E est un K-ev., une application $N : E \to \mathbb{R}^+$ est une norme ssi :*

$$\forall\, x \in E \quad N(x) = 0 \iff x = 0_E$$

$$\forall\, x \in E \quad \forall\, \lambda \in \mathbb{K} \quad N(\lambda x) = |\lambda| N(x) \text{ (homogénéité)}$$

$$\forall\, x, y \in E \quad N(x+y) \leqslant N(x) + N(y) \text{ (inégalité triangulaire).}$$

Un espace vectoriel normé est un couple (E, N), où N est une norme sur E.

EXEMPLE 6-2.2 Sur $\mathbb{K}^n$, on utilise souvent les normes "usuelles" suivantes : si $x = (x_1, \dots, x_n) \in \mathbb{K}^n$

- $\|x\|_1 = \sum_{i=1}^{n} |x_i|$
- $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$
- $\|x\|_\infty = \sup\{|x_i|,\ i = 1..n\}$

On vérifie aisément qu'il s'agit effectivement de normes, le seul résultat non immédiat étant peut-être l'inégalité triangulaire pour $\| \quad \|_2$, connue sous le nom d'inégalité de Minkowski, et qu'on reverra ultérieurement dans le cadre des espaces vectoriels euclidiens et hermitiens.

EXEMPLE 6-2.3 Sur l'espace vectoriel $\mathcal{C}^0([a,b],\mathbb{K})$ des fonctions continues sur le segment $[a,b]$ à valeurs dans $\mathbb{K}$, les expressions suivantes définissent également des normes :

- $\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$
- $\|f\|_1 = \int_a^b |f(t)|dt$
- $\|f\|_2 = \left(\int_a^b |f(t)|dt\right)^{\frac{1}{2}}$

L'existence de $\|f\|_\infty$ provenant du fait qu'une fonction numérique continue sur $[a,b]$ est bornée, vérifier qu'on définit ainsi une norme est élémentaire. Les implications

$$\|f\|_i = 0 \Longrightarrow f = 0$$

pour $i = 1$ ou 2 utilisent la continuité de la fonction f (voir chapitre sur l'intégrale). Enfin l'inégalité triangulaire pour $\| \quad \|_2$ porte encore le nom de Minkowski.[4]

PROPOSITION 6-2.4 Si x et y sont deux vecteurs d'un espace normé $(E, \| \quad \|)$, on a

$$|\|x\| - \|y\|| \leqslant \|x-y\|$$

Démonstration : Il suffit d'écrire $x = (x-y) + y$ et d'appliquer l'inégalité triangulaire. ■

DÉFINITION 6-2.5 (Vecteur unitaire) *Un vecteur x de $(E, \| \quad \|)$ est dit* ***unitaire*** *ssi $\|x\| = 1$.*

Si $x \neq 0_E$, les vecteurs unitaires et colinéaires à x sont de la forme λx avec $|\lambda|\,\|x\| = 1$. Ils sont de la forme $\pm\frac{x}{\|x\|}$ dans le cas réel, $e^{i\theta}\frac{x}{\|x\|}$, avec θ réel dans le cas complexe.

DÉFINITION 6-2.6 (Distance) *Sur l'espace normé $(E, \| \quad \|)$, on appelle* ***distance associée à la norme*** *l'application :*

$$d : E \times E \to \mathbb{R}^+ \textit{ définie par } (x,y) \mapsto d(x,y) = \|x-y\|$$

Cette application vérifie les propriétés

[4]Sur l'espace E des fonctions **continues par morceaux** sur $[a,b]$, les expressions définissant $\| \quad \|_1$ et $\| \quad \|_2$ ont encore un sens. L'application $N : E \to \mathbb{R}^+$ associée possède les propriétés d'une norme à part la propriété $N(x) = 0 \Rightarrow x = 0_E$ qui n'est plus vérifiée. On dit alors que N est une **semi-norme** sur E.

$d(x,y) = 0 \Leftrightarrow x = y$ *(d "sépare" les points de E)*

$d(x,y) = d(y,x)$ *(symétrie)*

$d(x,z) \leqslant d(x,y) + d(y,z)$ *(inégalité triangulaire)*

Ces propriétés découlent immédiatement des propriétés de la norme. On peut remarquer de plus que d est invariante par translation (pour tout $a \in E$ on a $d(x+a, y+a) = d(x,y)$) et que $d(\lambda x, \lambda y) = |\lambda| d(x,y)$ pour $\lambda \in \mathbb{K}$.

6-2.2 Boules et sphères

DÉFINITION 6-2.7 (Boules et sphères) *Si $(E, \| \ \|)$ est un espace normé, si $a \in E$ et r est un réel strictement positif, on appelle boule ouverte (resp. boule fermée, resp. sphère) de centre a et de rayon r les sous-ensembles de E :*

$B(a,r[= \{x \in E \mid \|x-a\| < r\}$.

$B(a,r] = \{x \in E \mid \|x-a\| \leqslant r\}$.

$S(a,r) = \{x \in E \mid \|x-a\| = r\}$.

Dans le cas particulier $a = 0_E$ et $r = 1$, on parle des **boules et sphère unité**. Les homothéties et translations opèrent sur E et permettent de reconstituer toutes les boules et sphères si on connaît les boules et sphères unité. Par exemple, il est clair que

$$B(a,r] = a + rB(0_E, 1] \text{ et } B(a,r[= a + rB(0_E, 1[$$

DÉFINITION 6-2.8 *Si a et b sont deux éléments de E, on appelle segment $[a,b]$ l'ensemble*

$$[a,b] = \{x \in E \mid \exists\, t \in [0,1] \quad x = (1-t)a + tb\}$$

C'est l'ensemble des barycentres de a et b affectés de coefficients positifs.

DÉFINITION 6-2.9 *Une partie A non vide d'un $\mathbb{K}$-ev est dite* ***convexe*** *si et seulement si*

$$\forall\, a, b \in A \quad [a,b] \subset A$$

Par exemple, la propriété de la borne supérieure permet de montrer que les seules parties convexes de $\mathbb{R}$ sont les intervalles (fermés, ouverts ou semi-ouverts).

PROPOSITION 6-2.10 Les boules ouvertes ou fermées d'un espace vectoriel normé sont convexes.

Démonstration : Montrons le par exemple pour une boule fermée $B = B(a,r]$. Si $x, y \in B$ et $t \in [0,1]$, on a

$$\begin{aligned}\|tx + (1-t)y - a\| &= \|t(x-a) + (1-t)(y-a)\| \\ &\leqslant t\,\|x-a\| + (1-t)\,\|y-a\| \leqslant r\end{aligned}$$

ce qui montre que $[x,y] \subset B$. ■

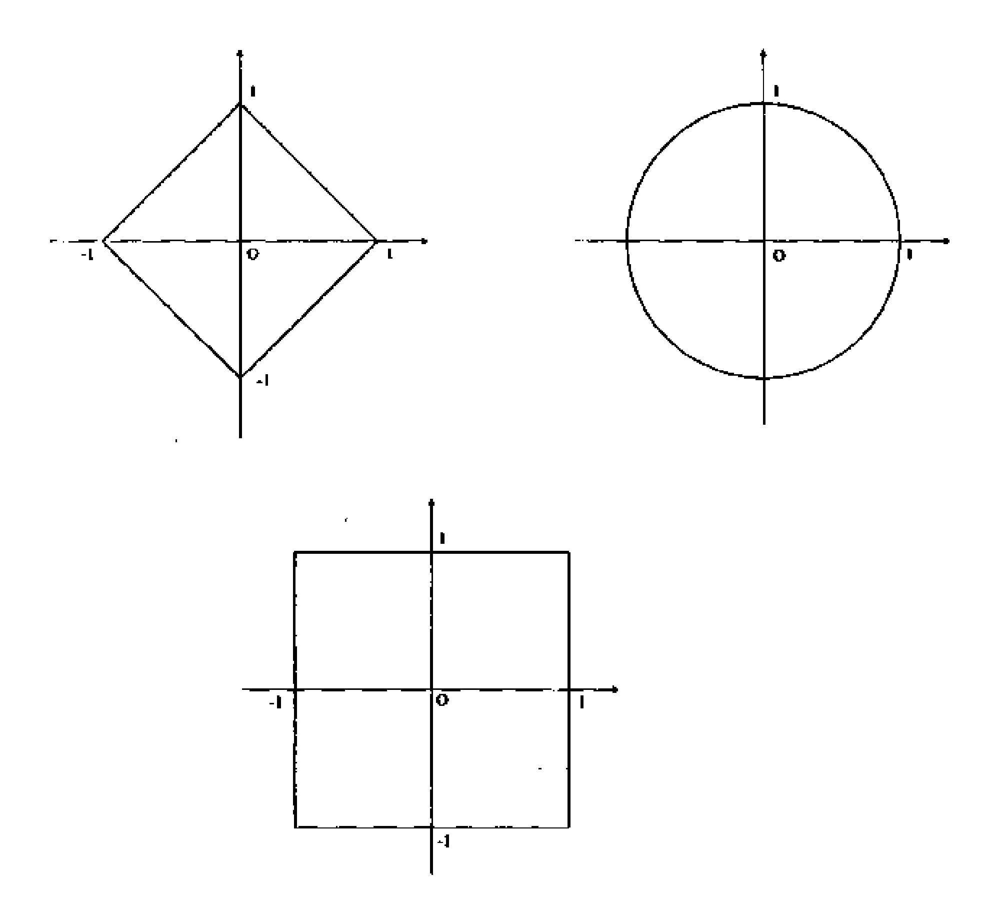


Figure **6.1** – *Boules unités de* $\mathbb{R}^2$ *pour* $\| \ \|_1$, $\| \ \|_2$ *et* $\| \ \|_\infty$

Dans l'espace $\mathbf{R}^2$ muni des normes "usuelles", les boules unités sont délimitées comme indiqué dans la figure **6.1**. On notera les inclusions

$$B_1(0,1[\subset B_2(0,1[\subset B_\infty(0,1[\subset B_1(0,\sqrt{2}[$$

Par homothéthie et translation, on constate sur cet exemple qu'une boule de centre a pour une des normes contient une boule de même centre pour toute autre de ces trois normes.

EXERCICE 6-2.11 Si $(E, \| \ \|)$ est un espace normé non réduit à $\{0_E\}$, montrer que

$$B(a,r] \subset B(a',r'] \iff r + \|a - a'\| \leqslant r'$$

En déduire que, dans un espace normé, l'égalité de deux boules équivaut à l'égalité des centres et des rayons.

6-2.3 Parties bornées, diamètre

DÉFINITION 6-2.12 (Partie bornée) *Une partie A d'un ev normé $(E, \| \ \|)$ est bornée ssi l'ensemble des normes $\{\|x\|,\ x \in A\}$ est majoré dans $\mathbf{R}^+$.*

Les propriétés suivantes découlent immédiatement de la définition :

- Toute partie contenue dans une partie bornée est bornée.
- Un réunion d'une famille **finie** de parties bornées est bornée.
- Toute boule est bornée. Une partie est bornée ssi elle est incluse dans une boule.

L'exemple (dans $(\mathbb{R}, |\,|)$) $\mathbb{R} = \bigcup_{n \in \mathbb{N}} [-n, n]$ montre qu'une union quelconque de parties bornées n'est pas bornée en général.

DÉFINITION 6-2.13 (Diamètre) *Si A est une partie bornée non vide de $(E, \|\ \|)$, on définit le* ***diamètre*** *de A par :*

$$\delta(A) = \sup\{\|x - y\|\,,\ x, y \in A\}$$

On remarquera que cette borne supérieure n'est pas atteinte en général par un couple (x, y) de $A \times A$.

EXERCICE 6-2.14 Montrer que, si $E \neq \{0_E\}$,

$$\delta\left(B(a,r[\right) = \delta\left(B(a,r]\right) = \delta\left(S(a,r)\right) = 2r$$

Si A est une partie bornée non vide, $a \in E$ et $\lambda \in \mathbb{K}$, comparer $\delta(A)$, $\delta(a + A)$, et $\delta(\lambda A)$. Si A et B sont bornées, que peut-on dire de $A + B = \{a + b,\ a \in A \text{ et } b \in B\}$?

DÉFINITION 6-2.15 *Si X est un ensemble non vide, une application f définie sur X à valeurs dans $(E, \|\ \|)$ est dite bornée ssi $f(X)$ est une partie bornée de E.*

THÉORÈME 6-2.16 (Norme de la convergence uniforme) L'ensemble des applications bornées de X dans E est un sous-espace vectoriel de E^X. On le note $\mathcal{B}(X, E)$. L'application $\|\ \|_\infty$ de $\mathcal{B}(X, E)$ dans $\mathbb{R}^+$ définie par

$$\|f\|_\infty = \sup\{\|f(x)\|\,,\ x \in X\}$$

est une norme sur $\mathcal{B}(X, E)$, appelée norme de la convergence uniforme sur X

Démonstration : Exercice. ■

En particulier, lorsque $X = \mathbb{N}$, on parle d'espace des suites bornées à valeurs dans E, et la norme est alors définie par $\|u\|_\infty = \sup_{n \in \mathbb{N}} \|u_n\|$ si u est la suite de terme général u_n. L'espace $\mathcal{B}(\mathbb{N}, E)$ est alors noté $l^\infty(\mathbb{N}, E)$ ou plus simplement $l^\infty(E)$.

6-2.4 Distance d'un point à une partie

DÉFINITION 6-2.17 *Si $\emptyset \neq A \subset E$ et $x \in E$, on définit la distance de x à A par :*

$$d(x, A) = \inf\{d(x,y),\ y \in A\}$$

Dans $(\mathbb{R}, |\ |)$, l'exemple $d(0,]0,1[) = 0$ montre qu'en général la distance d'un point à une partie n'est pas atteinte par un point de cette partie. On remarquera aussi que

$$d(x, A) = 0 \text{ n'est pas équivalent à } x \in A$$

Par contre, si $\alpha > 0$, on a l'équivalence :

$$d(x, A) < \alpha \iff B(a, \alpha[\cap A \neq \emptyset$$

Il faut se convaincre de la nécessité de travailler avec la boule ouverte et l'inégalité stricte.

THÉORÈME 6-2.18 (Inégalité triangulaire) **Si $\emptyset \neq A \subset E$ et $x, y \in E$ on a**

$$|d(x, A) - d(y, A)| \leqslant \|x - y\|$$

Démonstration : Si $z \in A$, on a $\|y - z\| \geqslant \|x - z\| - \|x - y\|$ par l'inégalité triangulaire. On en déduit, par définition de la borne inférieure, l'inégalité

$$\|y - z\| \geqslant d(x, A) - \|x - y\|$$

puis $d(y, A) \geqslant d(x, A) - \|x - y\|$ ce qui démontre l'inégalité souhaitée, puisque x et y y jouent des rôles symétriques. ■

6-2.5 Norme induite sur un sev. Distance induite sur une partie

Si $(E, \|\ \|)$ est un ev normé et si F est un sev de E, la restriction de la norme à F

$$\|\ \|_F : F \to \mathbb{R}^+ \quad x \mapsto \|x\|$$

est évidemment une norme sur F appelée norme induite sur F par $\|\ \|$. L'espace $(F, \|\ \|_F)$ est appelé sous-espace normé de $(E, \|\ \|)$. S'il n'y a pas d'ambiguïté, on commet un abus d'écriture en notant encore $\|\ \|$ la norme $\|\ \|_F$.

De même, si $X \subset E$ est une partie non vide de E, l'application d_X

$$d_X : X \times X \to \mathbb{R}^+ \quad (x, y) \mapsto d_X(x, y) = d(x, y) = \|x - y\|$$

restriction de la distance à X est appelée distance induite par la norme $\|\ \|$ sur X. On la notera souvent encore d, par abus d'écriture, et le couple (X, d) est appelé **espace métrique**. On pourrait définir dans X une notion de boule (ou de sphère) de centre $a \in X$, ensemble des points de X situés à une distance de a inférieure à un réel $r > 0$ donné. Il ne s'agit pas d'autre chose que des traces sur X des boules (ou sphères) dans E, de centre a et de rayon r. On remarquera cependant que, contrairement à ce qui se passe dans un espace normé $E \neq \{0_E\}$, une sphère de X peut être vide, et que deux boules peuvent être égales sans avoir même centre et même rayon.

6-2.6 Produit d'une famille finie d'espaces normés

DÉFINITION 6-2.19 (Produit d'espaces normés) *Si $(E_i, N_i)_{1\leqslant i\leqslant p}$ est une famille de p $\mathbb{K}$-ev normés le $\mathbb{K}$-ev $E = \prod_{1\leqslant i\leqslant p} E_i$ peut être muni d'une norme "naturelle" qu'on notera, s'il n'y a pas d'ambiguïté, N^∞ définie par*

$$\forall\, x = (x_1, \dots, x_p) \in E \quad N^\infty(x) = \sup\{N_i(x_i),\ 1 \leqslant i \leqslant p\}$$

L'espace (E, N^∞) est alors appelé ***espace produit*** *des espaces vectoriels normés $(E_i, N_i)_{1\leqslant i\leqslant p}$.*

On vérifie sans peine qu'il s'agit là bien d'une norme sur E. La distance associée sur E sera notée d_∞. On notera en particulier que, si $x = (x_1, \dots, x_p)$ et $y = (y_1, \dots, y_p)$ sont deux éléments de E, on a les inégalités :

$$\forall\, i \quad d_i(x_i, y_i) = N_i(x_i - y_i) \leqslant d_\infty(x, y) = N^\infty(x - y) \leqslant \sum_{j=1}^{p} N_j(x_j - y_j)$$

On remarquera également que l'espace normé $(\mathbb{K}^p, \|\ \|_\infty)$ est en fait le produit de p espaces normés $(\mathbb{K}, |\ |)$.

6-3 Suites d'un espace normé : convergence, valeurs d'adhérence

On généralise dans cette section les définitions et résultats donnés dans le cas des suites numériques dans la section **6-1**.

6-3.1 Convergence d'une suite

6-3.1.1 Suites d'un espace normé

On appelle suite de l'espace normé $(E, \|\ \|)$, toute application

$$\mathbb{N} \ni n \mapsto u(n) = u_n \in E$$

Par abus de langage, on parlera encore de suite lorsque l'application u n'est définie qu'à partir d'un certain rang. L'espace $E^{\mathbb{N}}$ des suites à valeurs dans E est un $\mathbb{K}$-ev. On sait que l'espace $l^\infty(E)$ des suites bornées en est un sev, qui est normé par la norme de la convergence uniforme.

6-3.1.2 Suites convergentes

DÉFINITION 6-3.1 (Suite convergente) *Une suite $(u_n)_{n\in\mathbb{N}}$ de vecteurs de l'espace normé $(E, \|\ \|)$ est convergente si et seulement s'il existe $l \in E$ vérifiant*

$$\forall\, \varepsilon > 0 \quad \exists\, N_\varepsilon \in \mathbb{N} \quad \forall\, n \geqslant N_\varepsilon \quad \|u_n - l\| \leqslant \varepsilon$$

ce qui revient à dire que la suite réelle (positive) $\|u_n - l\|$ est convergente vers 0. *Une suite non convergente est dite divergente.*

On remarquera donc que la convergence de u_n vers l équivaut à la convergence vers 0_E de la suite $u_n - l$. Comme dans le cas des suites complexes, les résultats suivants découlent directement de la définition :

THÉORÈME 6-3.2 (Unicité de la limite) Si une suite u est convergente, le vecteur l intervenant dans la définition est unique. On le note $\lim\limits_{n\to+\infty} u_n$. Toute suite convergente est bornée (la réciproque étant évidemment fausse).

Démonstration : Si l et l' répondent à la définition, le réel positif $\|l - l'\|$ vérifie

$$0 \leqslant \|l - l'\| \leqslant \|u_n - l\| + \|u_n - l'\|$$

La suite majorante tendant vers 0, on a forcément $\|l - l'\| = 0$ ce qui entraîne $l = l'$. De plus, si la suite converge vers l, on a, à partir d'un certain rang $\|u_n\| \leqslant 1 + \|l\|$, ce qui montre que la suite est bornée. ■

THÉORÈME 6-3.3 (Opérations sur les limites) Si $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ sont deux suites de $(E, \| \|)$ qui convergent respectivement vers l et l' et si $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ est une suite de $\mathbb{K}$ convergente vers $a \in \mathbb{K}$, la suite $u + \alpha v$ est convergente et

$$\lim_{n\to+\infty} (u_n + \alpha_n v_n) = l + al'$$

La suite réelle $(\|u_n\|)_{n\in\mathbb{N}}$ est convergente et

$$\lim_{n\to+\infty} \|u_n\| = \|l\|$$

En particulier, l'ensemble $C(E)$ des suites convergentes est un sous-espace vectoriel de l'espace normé $(l^\infty(E), \| \|_\infty)$ et l'application $u \mapsto \lim\limits_{n\to+\infty} u_n$ est linéaire de $C(E)$ dans E. Si la suite u est convergente on a

$$\left\|\lim_{n\to+\infty} u_n\right\| \leqslant \|u\|_\infty$$

Démonstration : Comme dans le cas des suites complexes, il suffit d'appliquer l'inégalité triangulaire

$$\begin{aligned}\|u_n + \alpha_n v_n - l - al'\| &= \|(u_n - l) + \alpha_n(v_n - l') + (\alpha_n - a)l'\| \\ &\leqslant \|u_n - l\| + |\alpha_n|\,\|v_n - l'\| + |\alpha_n - a|\,\|l'\|\end{aligned}$$

et de "couper les ε en 3". La deuxième assertion est conséquence de l'inégalité

$$|\|u_n\| - \|l\|| \leqslant \|u_n - l\|$$

Enfin, pour tout n on a $\|u_n\| \leqslant \|u\|_\infty$, ce qui donne la dernière inégalité. ■

THÉORÈME 6-3.4 (Suites dans un espace produit) Si (E, N^∞) est un produit d'espaces normés $\prod_{1\leqslant i\leqslant p} (E_i, N_i)$, une suite u de vecteurs de E, de terme général $u_n = (u_n^1, \dots, u_n^p)$ est convergente vers $l = (l_1, \dots, l_p) \in E$ ssi pour tout $i \in \{1, \dots, p\}$ la suite $(u_n^i)_{n\in\mathbb{N}}$ converge vers l_i dans (E_i, N_i). La convergence dans l'espace produit est donc la convergence "coordonnée par coordonnée".

Démonstration : C'est une conséquence immédiate des inégalités :

$$\forall n \in \mathbb{N} \quad \forall i \qquad N_i(u_n^i - l_i) \leqslant N^\infty(u_n - l) \leqslant \sum_{j=1}^{p} N_j(u_n^j - l_j)$$

déja signalées plus haut. ■

6-3.2 Comparaison de deux normes

Dans $\mathbb{K}^p$, dire qu'une suite converge vers une limite l pour $\| \|_\infty$ équivaut à dire qu'elle converge vers la même limite pour $\| \|_1$ (ou pour $\| \|_2$). Le fait de changer une des normes usuelles par une autre n'a aucune incidence sur la notion de suite convergente. Par contre, dans l'espace $\mathcal{C}^0([0,1], \mathbb{R})$, ce n'est plus le cas : La suite de fonctions $(f_n)_{n\in\mathbb{N}-\{0,1\}}$ définies par

$$f_n(x) = \begin{cases} nx & \text{si } 0 \leqslant x \leqslant 1/n \\ 2 - nx & \text{si } 1/n \leqslant x \leqslant 2/n \\ 0 & \text{si } 2/n \leqslant x \leqslant 1 \end{cases}$$

converge vers la fonction nulle pour $\| \|_1$ car $\|f_n\|_1 = \frac{1}{n}$ mais pas pour $\| \|_\infty$ car $\|f_n\|_\infty = 1$. On est donc amené à comparer la convergence vers 0_E de suites de vecteurs de E pour le choix de différentes normes.

THÉORÈME 6-3.5 Soit E un $\mathbb{K}$-ev et N et N' deux normes sur E. Pour que toute suite de E tendant vers 0_E pour N tende vers 0_E pour N', il faut et il suffit qu'il existe une constante $\alpha > 0$ telle que

$$\forall x \in E \quad N'(x) \leqslant \alpha N(x)$$

Démonstration : La condition est évidemment suffisante puisque, si α existe, pour toute suite u de E on a $0 \leqslant N'(u_n) \leqslant \alpha N(u_n)$. Réciproquement, si on ne peut trouver de valeur de α satisfaisant la propriété de majoration, on a

$$\forall n \in \mathbb{N}^* \quad \exists\, x_n \in E \quad N'(x_n) > nN(x_n)$$

Comme $N'(x_n) > 0$ on a $x_n \neq 0_E$. La suite $u_n = \dfrac{1}{n}\dfrac{x_n}{N(x_n)}$ converge évidemment vers 0_E pour N (puisque $N(u_n) = \dfrac{1}{n}$) mais pas pour N' puisque $N'(u_n) > 1$. ■

Par exemple, dans l'espace $E = C^0([a,b],\mathbb{R})$, la convergence pour la norme $\| \|_\infty$ entraîne la convergence pour $\| \|_1$ puisque, pour $f \in E$

$$\|f\|_1 = \int_0^1 |f(t)|dt \leqslant \sup_{t\in[0,1]} |f(t)| = \|f\|_\infty$$

(on dit que la convergence uniforme entraîne la convergence "en moyenne"). Par contre la réciproque est fausse, comme le montre le contre exemple précédent.

Dans $\mathbb{K}^p$, les notions de suites convergentes pour les trois normes "usuelles" sont équivalentes puisque

$$\forall\, x \in \mathbb{K}^p \quad \|x\|_\infty \leqslant \|x\|_2 \leqslant \|x\|_1 \leqslant p\,\|x\|_\infty$$

DÉFINITION 6-3.6 (Normes équivalentes) *Si N et N' sont deux normes sur E, on dit que N est* ***équivalente*** *à N' ssi il existe deux réels α et β strictement positifs tels que*

$$\forall\, x \in E \quad \alpha N(x) \leqslant N'(x) \leqslant \beta N(x)$$

On écrira en abrégé $\alpha N \leqslant N' \leqslant \beta N$.

THÉORÈME 6-3.7 La relation "être équivalente à" est une relation d'équivalence sur l'ensembles des normes sur le $\mathbb{K}$-ev E. Deux normes sur E sont équivalentes ssi elles définissent les mêmes suites convergentes (vers 0_E, donc par translation vers un élément l quelconque dans E).

Démonstration : Exercice. ■

EXEMPLE 6-3.8 Sur un produit d'espaces normés $E = \prod_{1\leqslant i\leqslant p} (E_i, N_i)$ on a défini la norme N^∞ qui définit la convergence "coordonnée par coordonnée". On vérifie aisément que les expressions

$$N^1(x_1,\dots,x_p) = \sum_{i=1}^{p} N_i(x_i) \quad \text{et} \quad N^2(x_1,\dots,x_p) = \sqrt{\sum_{i=1}^{p} (N_i(x_i))^2}$$

définissent sur E deux normes équivalentes à N^∞.

On peut donner une définition plus "géométrique" de l'équivalence des normes (donnée pour les boules de centre 0_E, mais valable pour les boules de centre quelconque, par translation) :

THÉORÈME 6-3.9 Deux normes sur un $\mathbb{K}$-ev sont équivalentes ssi toute boule (ouverte, par exemple) de centre 0_E pour l'une de ces normes contient une boule (ouverte) de centre 0_E pour l'autre norme.

Démonstration : Si on a un encadrement $\alpha N \leqslant N' \leqslant \beta N$, il est clair que pour $r > 0$

$$B_{N'}(0_E, \alpha r[\subset B_N(0_E, r[\quad \text{et} \quad B_N(0_E, \frac{r}{\beta}[\subset B_{N'}(0_E, r[$$

Réciproquement, s'il existe $a > 0$ avec $B_N(0_E, a[\subset B_{N'}(0_E, 1[$, si x non nul est dans E, le vecteur $y = \frac{a}{2}\frac{x}{N(x)}$ est dans $B_N(0_E, a[$ et vérifie donc $N'(y) \leqslant 1$, ce qui donne $N'(x) \leqslant \frac{2}{a}N(x)$. Comme cette inégalité est encore vraie si $x = 0_E$, on en déduit l'existence d'un réel $\alpha > 0$ tel que $N' \leqslant \alpha N$. Comme les rôles joués par N et N' sont symétriques, l'équivalence de N et N' en découle. ■

EXERCICE 6-3.10 Montrer que les normes N et N' sur E sont équivalentes ssi les espaces normés (E, N) et (E, N') ont les mêmes parties bornées, les mêmes suites bornées.

6-3.3 Valeur d'adhérence d'une suite

Comme dans les cas réel et complexe on donne la définition et la caractérisation :

DÉFINITION 6-3.11 *Si $u = (u_n)_{n \in \mathbb{N}}$ est une suite de $(E, \| \,\|)$ et si $l \in E$, on dit que l est une* ***valeur d'adhérence*** *de la suite u si et seulement s'il existe une suite extraite de u qui converge vers l.*

Cette propriété est invariante lorsqu'on remplace la norme par une norme équivalente.

PROPOSITION 6-3.12 Un vecteur $l \in E$ est valeur d'adhérence de la suite u ssi, pour tout $\varepsilon > 0$, l'ensemble

$$\{n \in \mathbb{N} \mid u_n \in B(l, \varepsilon[\}$$

est infini.

Toute suite convergente ne possède qu'une valeur d'adhérence qui est sa limite. En conséquence, une suite ayant au moins deux valeurs d'adhérence distinctes ne peut converger. Comme pour le passage de $\mathbb{R}$ à $\mathbb{C}$, par extraction successives de suites pour les différentes coordonnées on démontre le

THÉORÈME 6-3.13 (Bolzano-Weierstrass)

Bolzano-Weierstrass Dans $\mathbb{K}^p$ muni d'une des normes usuelles, toute suite bornée possède au moins une valeur d'adhérence.

COROLLAIRE 6-3.14 Une suite bornée de $(\mathbb{K}^p, \| \,\|_\infty)$ converge ssi elle possède une unique valeur d'adhérence.

On prendra garde que ce théorème (ou son corollaire) ne se généralise pas à des espaces normés plus généraux (de dimension infinie). Par exemple dans l'espace $(\mathcal{B}([0,1], \mathbb{R}), \| \,\|_\infty)$ des fonctions numériques bornées sur $[0,1]$, muni de la norme de la convergence uniforme, la suite de fonctions $f_n = \mathbf{1}_{[0,\frac{1}{n}[}$ (fonction caractéristique de l'intervalle $[0, \frac{1}{n}[$) est bornée puisque $\|f_n\|_\infty = 1$. Pour $n \neq m$, on vérifie que $\|f_n - f_m\|_\infty = 1$, ce qui montre facilement qu'aucune suite extraite de la suite $(f_n)_{n \in \mathbb{N}}$ ne peut converger.

6-4 Topologie d'un espace normé

6-4.1 Voisinages d'un point

DÉFINITION 6-4.1 (Voisinage) *Si $(E, \| \|)$ est un espace normé et si $a \in E$, une partie $V \subset E$ est un* **voisinage** *de a ssi il existe un réel $r > 0$ tel que $B(a, r[\subset V$. On notera $\mathcal{V}(a)$ le sous-ensemble de l'ensemble des parties de E formé de tous les voisinages de a.*

Les propriétés suivantes découlent immédiatement de la définition :

- a appartient à tout voisinage de a.
- Toute partie contenant un voisinage de a est un voisinage de a.
- L'intersection d'une famille **finie** de voisinages de a est un voisinage de a.
- Par contre l'intersection d'une famille infinie de voisinages n'est pas, en général, un voisinage. Par exemple

$$\bigcap_{V \in \mathcal{V}(a)} V = \{a\}$$

puisque les boules de centre a sont des voisinages de a et que

$$\bigcap_{n \in \mathbb{N}^*} B(a, \frac{1}{n}[= \{x \in E \mid \|x - a\| = 0\} = \{a\}$$

- Si $a \neq b$ sont deux points de E, il existe des voisinages $V \in \mathcal{V}(a)$ et $V' \in \mathcal{V}(b)$ avec $V \cap V' = \emptyset$. On dit qu'un ev normé est **séparé**.

REMARQUE 6-4.2 Si N et N' sont deux normes équivalentes sur E, la notion de voisinage est la même dans les espaces (E, N) et (E, N'), à cause des inclusions de boules résultant de l'équivalence des normes.

Avec le vocabulaire que l'on vient d'introduire on a d'ailleurs :

PROPOSITION 6-4.3 Une suite $u = (u_n)_{n \in \mathbb{N}}$ de vecteurs de E converge vers $l \in E$ ssi

$$\forall V \in \mathcal{V}(l) \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad u_n \in V$$

REMARQUE 6-4.4 Dans le cas particulier de $(\mathbb{R}, | |)$, une partie est dite voisinage de a à gauche ssi elle contient un voisinage de la forme $]a - r, a]$, avec $r > 0$. Il ne s'agit pas d'un voisinage de a au sens de la définition générale donnée ici.

DÉFINITION 6-4.5 (Voisinage de l'infini) *On dit qu'une partie $V \subset (E, \| \|)$ est un voisinage de l'infini dans E ssi il existe un réel $M > 0$ tel que*

$$V \supset \{x \in E \mid \|x\| \geqslant M\}$$

Il est équivalent de dire que V contient le complémentaire d'une partie bornée de E. Les propriétés des voisinages énoncées plus haut se modifient de manière évidente. En particulier

$$\bigcap_{V \in \mathcal{V}(\infty)} V = \emptyset$$

6-4.2 Ouverts d'un espace normé

DÉFINITION 6-4.6 (Ouverts) *Une partie O d'un espace normé $(E, \| \|)$ est un ouvert ssi*

$$\forall x \in O \quad \exists \varepsilon > 0 \quad B(x, \varepsilon[\subset O$$

Il est équivalent de dire que O est voisinage de chacun de ses points. L'ensemble des ouverts de $(E, \| \|)$ est appelé **topologie** *de l'espace normé E. Les ouverts ne changent pas lorsqu'on remplace la norme par une norme équivalente.*

THÉORÈME 6-4.7 Les propriétés suivantes découlent immédiatement de la définition et des propriétés des voisinages : E et $\emptyset$ sont des ouverts de E. Une réunion d'une famille quelconque d'ouverts de E est un ouvert de E. Une intersection d'une famille finie d'ouverts de E est un ouvert de E.

On prendra garde qu'une intersection quelconque d'ouverts n'est pas un ouvert en général. Par exemple, dans $(\mathbb{R}, | |)$, l'intersections des ouverts $]-\frac{1}{n}, \frac{1}{n}[$ pour n parcourant $\mathbb{N}^*$ n'est pas ouvert.

PROPOSITION 6-4.8 Toute boule ouverte de E est un ouvert. Une partie non vide de E est un ouvert ssi elle est réunion d'une famille de boules ouvertes.

Démonstration : Si $O = B(a, r[$ est une boule ouverte et $x \in O$, on vérifie aisément, à l'aide de l'inégalité triangulaire, que

$$B\left(x, \frac{r - \|x\|}{2}\right[\subset O$$

$B(a, r[$ est donc bien voisinage de chacun de ses points. Toute boule ouverte étant un ouvert, une réunion de boules ouvertes est ouverte. Réciproquement, si O est un ouvert non vide, on a

$$\forall x \in O \quad \exists \varepsilon_x > 0 \quad B(x, \varepsilon_x[\subset O$$

O peut alors s'écrire comme union de boules ouvertes

$$O = \bigcup_{x \in O} B(x, \varepsilon_x[\quad \blacksquare$$

EXERCICE 6-4.9 Montrer qu'une boule fermée d'un espace normé non réduit à $\{0_E\}$ n'est jamais ouverte.

DÉFINITION 6-4.10 (Ouvert élémentaire) *Si $(E, N^\infty) = \prod_{1 \leqslant i \leqslant p} (E_i, N_i)$ est un produit d'espaces normés, on appelle ouvert élémentaire de E toute partie de E de la forme*

$$O = O_1 \times O_2 \times \cdots \times O_p$$

où O_i est un ouvert de E_i.

THÉORÈME 6-4.11 Tout ouvert élémentaire de E est un ouvert de (E, N^∞). Une partie de E est un ouvert de (E, N^∞) ssi elle est réunion d'ouverts élémentaires.

Démonstration Si $O = O_1 \times O_2 \times \cdots \times O_p$ est ouvert élémentaire, il n'y a rien à dire si l'un des O_i est vide. Sinon, si $x = (x_1, \ldots, x_p) \in O$, pour tout i on a $x_i \in O_i$ ouvert de E_i. Il existe donc $r_i > 0$ tel que $B_{N_i}(x_i, r_i[\subset O_i$. Si $r = \min(r_1, \ldots, r_p)$, il est clair que $B_{N_\infty}(x, r[\subset O$. Un ouvert élémentaire est donc ouvert, il en est de même d'une réunion d'ouverts élémentaires. Réciproquement, un ouvert de E est réunion de boules ouvertes pour N_∞ et il est clair, avec les notations précédentes que, si $a > 0$

$$B_{N_\infty}(x, a[= \prod_{i=1}^{p} B_{N_i}(x_i, a[$$

est ouvert élémentaire. ■

6-4.3 Fermés d'un espace normé

DÉFINITION 6-4.12 (Fermés d'un ev normé) *Une partie F d'un espace vectoriel normé $(E, \| \|)$ est un fermé ssi son complémentaire $F^c = E - F$ est un ouvert de E.*

Ici encore, les fermés ne changent pas si on remplace la norme par une norme équivalente.

THÉORÈME 6-4.13 Des propriétés des ouvert découlent immédiatement celles des fermés, par passage au complémentaire : E et $\emptyset$ sont des fermés de E. Une intersection d'une famille quelconque de fermés de E est un fermé de E. Une union d'une famille finie de fermés de E est un fermé de E.

EXERCICE 6-4.14 Montrer que toute boule fermée, toute sphère d'un espace normé est fermée. Montrer qu'une boule ouverte n'est jamais fermée si $E \neq \{0_E\}$.

6-4.4 Point adhérent, adhérence

DÉFINITION 6-4.15 (Point adhérent) *Si $A \subset (E, \| \|)$, un point $x \in E$ est dit adhérent à A ssi*

$$\forall V \in \mathcal{V}(x) \quad V \cap A \neq \emptyset$$

Il est équivalent de dire (à cause de la définition des voisinages) que toute boule ouverte centrée en x rencontre A, ou encore que tout ouvert contenant x rencontre A.

Comme on a vu que $B(x, \alpha[\cap A \neq \emptyset \Leftrightarrow d(x, A) < \alpha$, on a immédiatement la

PROPOSITION 6-4.16 Un point $x \in E$ est adhérent à une partie A (non vide) de E ssi $d(x, A) = 0$.

DÉFINITION 6-4.17 (Adhérence d'une partie) *Si A est une partie de $(E, \| \|)$, on appelle adhérence de A et on note $\overline{A}$ l'ensemble des points adhérents à A.*

THÉORÈME 6-4.18 Si $A \subset E$, $\overline{A}$ est un fermé. C'est le plus petit fermé (au sens de l'inclusion) contenant A.

Démonstration : Si $A = \emptyset$, il n'y a rien à dire. Sinon, il est d'abord clair sur la définition que $A \subset \overline{A}$. On montre ensuite que $\overline{A}$ est fermé, en étudiant son complémentaire. Si $x \notin \overline{A}$, il existe un $\alpha > 0$ tel que $B(x, \alpha[\cap A = \emptyset$. Si $y \in B(x, \alpha[$, cette boule est un voisinage de y (c'est un ouvert) qui ne rencontre pas A, et on a donc $y \notin \overline{A}$. $B(x, \alpha[$ est donc incluse dans le complémentaire de $\overline{A}$ qui est bien ouvert. Enfin, montrons que tout fermé F contenant A contient $\overline{A}$. Si $x \notin F$, le complémentaire F^c de F dans E est un ouvert qui contient x et qui ne rencontre pas A (puisque $F \supset A$). Donc $x \notin \overline{A}$. On a donc prouvé que $F^c \subset \overline{A}^c$, ce qui prouve l'inclusion de $\overline{A}$ dans F. ■

COROLLAIRE 6-4.19 Si O est un ouvert de E on a :

$$O \cap A \neq \emptyset \Leftrightarrow O \cap \overline{A} \neq \emptyset$$

COROLLAIRE 6-4.20 Une partie A de E est fermée si et seulement si

$$A = \overline{A}$$

REMARQUE 6-4.21 Le théorème précédent donne donc une autre définition de l'adhérence de A. Si A est une partie de E, il existe des parties fermées de E contenant A (par exemple E). On peut donc considérer

$$\overline{A} = \bigcap_{\substack{F \text{ fermé de } E \\ F \supset A}} F$$

qui est fermé comme intersection de fermés, et est par construction le plus petit fermé contenant A. C'est pour cette raison que l'adhérence d'une partie A est parfois appelée **fermeture** de A.

Les résultats suivants peuvent être démontrés en appliquant directement la définition des points adhérents. Les caractérisations données dans le paragraphe suivant en fourniront d'autres démonstrations :

EXERCICE 6-4.22 Montrer que $\overline{B(a, r[} = B(a, r]$ et que

$$\overline{\{x \in E \mid \|x - a\| > r\}} = B(a, r[^c$$

EXERCICE 6-4.23 Si $A \subset B$ montrer que $\overline{A} \subset \overline{B}$. Montrer que, si A et B sont des parties quelconques de E, on a

$$\overline{A \cup B} = \overline{A} \cup \overline{B} \quad \text{et} \quad \overline{A \cap B} \subset \overline{A} \cap \overline{B}$$

L'exercice précédent donne un exemple où l'inclusion est stricte. Un autre exemple classique dans $(\mathbb{R}, | |)$ est donné par $A = \mathbb{Q}$ et $B = \mathbb{R} - \mathbb{Q}$. Montrer qu'en général, si $(A_i)_{i \in I}$ est une famille quelconque de parties de E, il existe une relation d'inclusion entre $\overline{\cup A_i}$ et $\cup \overline{A_i}$. Etudier notamment, dans $\mathbb{R}$, le cas $A_n = [0, 1 - \frac{1}{n}[$ pour $n \in \mathbb{N}^*$.

EXERCICE 6-4.24 Si $u = (u_n)_{n\in\mathbb{N}}$ est une suite de points de E, on note

$$U_n = \{u_k \mid k \geqslant n\}$$

Montrer que l'ensemble des valeurs d'adhérence de la suite u est exactement $\bigcap_{n\in\mathbb{N}} \overline{U_n}$. En déduire que l'ensemble des valeurs d'adhérence d'une suite est un fermé (éventuellement vide) de E.

6-4.5 Caractérisation séquentielle de l'adhérence, des parties fermées

THÉORÈME 6-4.25 (Adhérence et suite convergente) Si A est une partie non vide de E, un point $x \in E$ est adhérent à A ssi il existe une suite de points de A qui converge vers x.

Démonstration : Si (a_n) est une suite de points de A convergente vers x, tout voisinage de x contient tous les termes de la suite à partir d'un certain rang, donc rencontre A, et par conséquent $x \in \overline{A}$. Réciproquement, si $x \in \overline{A}$, pour tout $n \in \mathbb{N}^*$ $A \cap B(x, \frac{1}{n}[\neq \emptyset$ et on peut donc trouver un point $a_n \in A$ avec $\|a_n - x\| < \frac{1}{n}$. La suite a_n est bien convergente vers x. ■

THÉORÈME 6-4.26 Une partie $A \subset E$ est fermée ssi toute suite convergente (dans $(E, \| \|)$) de points de A a sa limite qui appartient à A. Une partie fermée est donc, d'une certaine manière, "stable pour le passage à la limite". Il revient évidemment au même de dire que toutes les valeurs d'adhérence éventuelles d'une suite de points de A sont dans A.

Démonstration : Si A est fermée, on a $A = \overline{A}$. Si une suite de points de A converge, c'est forcément vers un point de $\overline{A}$ d'après le théorème précédent, donc vers un point de A. Réciproquement, si $l \in \overline{A}$, il existe une suite de points de A qui converge vers l. Si la propriété de l'énoncé est vérifiée, on a $l \in A$, donc $A = \overline{A}$ est fermé. ■

COROLLAIRE 6-4.27 Dans un espace normé, l'adhérence d'un sous-espace vectoriel est un sous-espace vectoriel.

Démonstration : Si F est un sev de E, on a $F \subset \overline{F} \neq \emptyset$. De plus si x et y sont deux points de $\overline{F}$, et si $\alpha \in \mathbb{K}$, il existe des suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ de points de F qui convergent respectivement vers x et y. La suite de F de terme général $x_n + \alpha y_n$ est donc convergente vers $x + \alpha y$ qui est donc bien dans $\overline{F}$. ■

Dans $(\mathbb{K}^p, \| \|_\infty)$, un sev est toujours fermé puisque défini, s'il est de dimension r, par un système de $p - r$ équations linéaires indépendantes de la forme

$$\alpha_1^j x_1 + \alpha_2^j x_2 + \cdots + \alpha_p^j x_p = 0 \quad \text{pour } 1 \leqslant j \leqslant p - r$$

Comme la convergence dans $\mathbb{K}^p$ est la convergence coordonnée par coordonnée, il est clair qu'une limite d'une suite convergente dont le terme général vérifie ces $p-r$ conditions les vérifie également (théorème sur les combinaisons linéaires de suites convergentes de $\mathbb{K}$). Par contre, dans des espaces normés plus généraux, on peut avoir des sev non fermés, pour lesquels l'inclusion $F \subsetneq \overline{F}$ est stricte. Voir par exemple :

EXERCICE 6-4.28 Dans $(C^0([a,b],\mathbb{R}), \| \|_1)$, montrer que

$$F = \{f \in C^0([a,b],\mathbb{R}) \mid f(0) = 0\}$$

est un sev vérifiant $\overline{F} = C^0([a,b],\mathbb{R})$.

6-4.6 Partie dense

DÉFINITION 6-4.29 (Partie dense) *Une partie $A \subset E$ est dite dense dans E ssi $\overline{A} = E$.*

Les caractérisations de l'adhérence qui précèdent donnent immédiatement :

THÉORÈME 6-4.30 Une partie A de E est dense dans E ssi tout point de E est limite d'une suite convergente de points de A. De manière équivalente, A est dense ssi tout ouvert non vide de E rencontre A.

EXEMPLE 6-4.31 Dans $\mathbb{R}$, une conséquence de l'existence de la partie entière d'un réel quelconque (elle même conséquence de la propriété de la borne supérieure) est que tout intervalle non réduit à un point contient un rationnel et un irrationnel, et donc une infinité de tels points. Comme un intervalle de ce type contient toujours un intervalle ouvert, cette propriété traduit exactement la **densité de $\mathbb{Q}$ et $\mathbb{R}-\mathbb{Q}$** dans $(\mathbb{R}, | |)$.

L'exemple précédent montre que l'intersection de deux parties denses n'est pas dense en général. Par contre on a :

EXERCICE 6-4.32 Une intersection d'un nombre fini d'**ouverts denses** est dense.

La densité de $\mathbb{Q}$ dans $\mathbb{R}$ peut être considérée comme un cas particulier du théorème suivant :

THÉORÈME 6-4.33 Un sous-groupe additif de $\mathbb{R}$ est soit de la forme $\alpha\mathbb{Z}$, soit dense dans $\mathbb{R}$.

Démonstration : On remarquera donc que les sous-groupes additifs fermés de $\mathbb{R}$ sont, soit de la forme $\alpha\mathbb{Z}$, soit égaux à $\mathbb{R}$. Si H est un sous-groupe non réduit à $\{0\}$, on a $H \cap \mathbb{R}^{*+} \neq \emptyset$. Cette partie étant minorée par 0, elle possède une borne inférieure. Si cette borne est > 0, notons la α, et montrons que $H = \alpha\mathbb{Z}$. On commence par montrer que $\alpha \in H$. Si ce n'est pas le cas, la définition de la borne supérieure montre l'existence d'un élément $x \in H \cap]\alpha, 2\alpha[$, puis d'un élément $y \in H \cap]\alpha, x[$. Mais alors $z = x - y$ est dans H (groupe) et vérifie $0 < z < \alpha$, ce qui est contradictoire avec la définition de α.

Donc $\alpha \in H$ et, H étant un groupe, $\alpha\mathbb{Z} \subset H$. L'inclusion inverse provient du fait que tout $x \in H$ peut s'écrire de manière unique sous la forme

$$x = p\alpha + r \quad \text{avec } p \in \mathbb{Z} \text{ et } 0 \leqslant r < \alpha$$

(c'est toujours l'existence de la partie entière). $r = x - p\alpha$ ne peut être dans H que s'il est nul. Enfin, si la borne inférieure de $H \cap \mathbb{R}^{*+}$ est nulle, on montre que H est dense. Si $]a, b[$ est un intervalle ouvert de $\mathbb{R}$, il existe un élément $x \in H \cap]0, b-a[$. Si p est la partie entière de $\frac{a}{x}$, on a

$$px \leqslant a < (p+1)x < px + b - a \leqslant b$$

ce qui montre que $(p+1)x \in H \cap]a, b[$. H est bien dense dans $\mathbb{R}$. ■

EXERCICE 6-4.34 Montrer que tout réel est limite d'une suite de la forme

$$u_n = p_n + q_n\sqrt{2}$$

où p_n et q_n sont des suites d'entiers relatifs.

EXERCICE 6-4.35 Si $\theta \notin \pi\mathbb{Q}$, montrer que $\theta\mathbb{Z}+2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$. En déduire que l'ensemble des valeurs d'adhérence de la suite $\left(e^{in\theta}\right)_{n\in\mathbb{N}}$ est le cercle unité du plan complexe.

EXERCICE 6-4.36 Dans un espace normé, montrer qu'un hyperplan est fermé ou dense. L'exercice **6-4**.28 donne un exemple d'hyperplan dense.

La propriété de densité admet la généralisation suivante, en relation avec la notion de topologie induite sur une partie d'un espace normé qui sera abordée à la fin de cette section :

DÉFINITION 6-4.37 (Densité relative) *Si $(E, \| \|)$ est un espace normé et si $A \subset B$ sont deux parties de E, on dit que A est dense dans B ssi $B \subset \overline{A}$. Ceci équivaut encore à dire que tout point de B est limite d'une suite convergente de points de A.*

THÉORÈME 6-4.38 Si $A \subset B \subset C$ sont trois parties de E, si A est dense dans B et B est dense dans C, alors A est dense dans C.

Démonstration : $B \subset \overline{A}$ implique $\overline{B} \subset \overline{A}$ (l'adhérence de B est le plus petit fermé contenant B). Or par hypothèse $C \subset \overline{B}$, donc on a bien $C \subset \overline{A}$: A est dense dans C. Ce raisonnement traduit l'idée fort intuitive suivante, mais peut-être délicate à formaliser : si tout point de B est limite d'une suite de points de A et si tout point de C est limite d'une suite de points de B, alors tout point de C peut être approché par une suite convergente de points de A. ■

Terminons enfin par un exemple classique de partie dense dans un evn. :

EXEMPLE 6-4.39 L'espace des matrices carrées $M_p(\mathbb{K})$ est identifié à l'espace $\mathbb{K}^{p^2}$, et est muni d'une des trois normes usuelles, par exemple

$$\|(m_{i,j})\| = \max\{|m_{i,j}|, \quad 1 \leqslant i,j \leqslant p\}$$

La convergence d'une suite de matrices pour cette norme est donc la convergence "coefficient par coefficient". Dans cet espace, $GL_p(\mathbb{K})$, ensemble des matrices inversibles est un **ouvert dense**. La densité est, par exemple, conséquence du fait que le spectre d'une matrice $A \in M_p(\mathbb{K})$ étant fini, pour n suffisamment grand la matrice $A_n = A - \frac{1}{n}I_p$ est inversible et $\lim_{n\to+\infty} A_n = A$. On peut d'ailleurs remarquer que ce résultat ne dépend pas de la norme choisie, ce qui n'a rien d'étonnant puisqu'on verra ultérieurement que toutes les normes sur $M_p(\mathbb{K})$ sont équivalentes. Il est facile (exercice), en utilisant par exemple le déterminant, de montrer que l'ensemble des matrices non inversibles est fermé dans $(M_p(\mathbb{K}), \|\ \|_\infty)$.

6-4.7 Intérieur, frontière

DÉFINITION 6-4.40 (Intérieur) *Si $A \subset E$, un point $a \in A$ est dit intérieur à A ssi A est un voisinage de a, ce qui équivaut à dire qu'il existe une boule ouverte centrée en a incluse dans A. On appelle intérieur de A et on note $\overset{\circ}{A}$ l'ensemble des points intérieurs à A.*

THÉORÈME 6-4.41 L'intérieur de A est le plus grand ouvert inclus dans A. C'est aussi la réunion de tous les ouverts inclus dans A. A est ouvert ssi $A = \overset{\circ}{A}$

Démonstration : $\emptyset$ est un ouvert inclus dans A. Si $\Omega = \bigcup_{\substack{\omega \text{ ouvert} \\ \omega\subset A}} \omega$, Ω est clairement le plus grand ouvert inclus dans A. Tout point de Ω est dans un ouvert $\omega \subset A$ et est donc intérieur à A. Réciproquement, si a est intérieur à A, il existe une boule ouverte ω centrée en a et incluse dans A. Donc $a \in \Omega$. ■

EXERCICE 6-4.42 Montrer que l'intérieur de $B(a, r]$ est $B(a, r[$.

COROLLAIRE 6-4.43 Le complémentaire de l'intérieur de A est égal à l'adhérence du complémentaire de A. De même le complémentaire de l'adhérence de A est égal à l'intérieur du complémentaire de A.

La deuxième assertion est conséquence de la première, en remplaçant A par son complémentaire. Pour démontrer que

$$\left(\overset{\circ}{A}\right)^c = \overline{A^c}$$

il suffit de remarquer que le complémentaire d'un ouvert inclus dans A est un fermé contenant le complémentaire de A. Si l'on part du plus grand ouvert inclus dans A, on obtient le plus petit fermé contenant A^c. ■

EXERCICE 6-4.44 Le passage au complémentaire permet d'obtenir les propriétés de l'intérieur à partir de celles de l'adhérence. On a en particulier :

$$\overset{\circ}{\widehat{A \cap B}} = \overset{\circ}{A} \cap \overset{\circ}{B} \quad \text{et} \quad \overset{\circ}{\widehat{A \cup B}} \supset \overset{\circ}{A} \cup \overset{\circ}{B}$$

DÉFINITION 6-4.45 (Frontière) *On appelle frontière de la partie $A \subset E$ et on note $Fr(A)$ le sous-ensemble de E égal à*

$$Fr(A) = \overline{A} - \overset{\circ}{A} = \overline{A} \cap \overline{A^c}$$

Les points de $Fr(A)$ sont appelés points-frontière de A.

La définition des adhérences de A et de son complémentaire montrent qu'un point x est dans la frontière de A ssi tout voisinage de x rencontre à la fois les parties A et A^c.

EXERCICE 6-4.46 Montrer que

$$\mathrm{Fr}(B(a,r[) = \mathrm{Fr}(B(a,r]) = S(a,r)$$
$$A \text{ est ouvert} \Leftrightarrow A \cap \mathrm{Fr}(A) = \emptyset$$
$$A \text{ est fermé} \Leftrightarrow A \supset \mathrm{Fr}(A)$$

6-4.8 Topologie induite sur une partie

DÉFINITION 6-4.47 (Voisinage relatif) *Si A est une partie d'un espace normé, $a \in A$ et $V \subset A$, on dit que V est un voisinage de a relatif à A ssi il existe un réel $r > 0$ tel que $B(a,r[\cap A \subset V$. Il revient au même de dire qu'il existe une boule ouverte de centre a pour la distance induite sur A incluse dans V.*

Les propriétés des voisinages restent vérifiées pour l'essentiel : toute partie de A contenant un voisinage de a relatif à A est encore un tel voisinage, une intersection d'un nombre fini de voisinages relatifs de a est encore un voisinage relatif, l'intersection de tous les voisinages relatifs de a est égale à $\{a\}$, deux points distincts de A possèdent des voisinages relatifs disjoints.

THÉORÈME 6-4.48 Une partie V de A est un voisinage relatif de $a \in A$ ssi c'est l'intersection avec A d'un voisinage de a (dans l'espace normé E).

> *Démonstration* : Si V est un voisinage relatif de a, il existe un réel $r > 0$ tel que l'on ait $B(a,r[\cap A \subset V$. On a alors évidemment $V = (V \cup B(a,r[) \cap A$, qui est bien trace sur A d'un voisinage $V' = V \cup B(a,r[$ de a dans $(E, \| \ \|)$. Réciproquement, si $V = V' \cap A$ avec $V' \in \mathcal{V}(a)$, il existe $r > 0$ avec $B(a,r[\subset V'$, donc $B(a,r[\cap A \subset V$ et V est un voisinage relatif de a dans A. ■

Par exemple, dans $(\mathbb{R}, | \ |)$, $]0,1]$ est un voisinage de 1 relatif à $A =]-1,1] \cup]2,4[$, puisque c'est la trace sur A du voisinage de 1 égal (par exemple) à $]0,2[$. Un voisinage de a relatif à une partie A n'est donc pas en général un voisinage de a dans E.

DÉFINITION 6-4.49 (Ouvert relatif) *Si A est une partie non vide de E, une partie ω de A est un ouvert relatif de A si ω est voisinage relatif à A de chacun de ses points.*

Comme dans le cas des voisinages, on a la caractérisation à l'aide des ouverts de E :

THÉORÈME 6-4.50 Une partie ω de A est un ouvert relatif de A ssi c'est l'intersection avec A d'un ouvert de l'espace normé E.

Démonstration : La caractérisation des voisinages relatifs donnée dans le théorème précédent montre immédiatement que si la partie $\omega = \Omega \cap A$ est trace sur A d'un ouvert Ω de E, ω est ouvert relatif de A (puisque Ω est voisinage de chacun de ses points). Réciproquement, si ω est un ouvert relatif de A, pour tout $x \in \omega$, il existe $r_x > 0$ avec $B(x, r_x[\cap A \subset \omega$. On a alors

$$\omega = \left(\bigcup_{x \in \omega} B(x, r_x[\right) \bigcap A$$

est bien trace sur A d'un ouvert (union de boules ouvertes) de E. ∎

L'ensemble des ouverts relatifs de A est appelé **topologie induite sur A** par la topologie de E. On vérifie que cette famille de parties de A vérifie les trois propriétés déja vérifiées par la topologie de E : $\emptyset$ et A sont des ouverts, une union quelconques d'ouverts de A est un ouvert de A, une intersection d'une famille finie d'ouverts de A est un ouvert de A.

Il est clair qu'un ouvert de E inclus dans A est un ouvert de A (il est égal à sa propre trace sur A) mais un ouvert de A n'est pas en général ouvert de E. Par exemple $\{0\}$ est ouvert de $A = \{0\} \cup [1, 2] \subset \mathbf{R}$.

DÉFINITION 6-4.51 (Fermé relatif) *Une partie $\phi \subset A$ est un fermé (relatif) de A ssi son complémentaire dans A est un ouvert relatif de A.*

Si $A - \phi = \Omega \cap A$ avec Ω ouvert de E, on a évidemment $\phi = \Omega^c \cap A$ (le complémentaire de Ω est calculé dans E). On a donc immédiatement la caractérisation des fermés relatifs :

THÉORÈME 6-4.52 Une partie ϕ de A est un fermé relatif de A ssi c'est la trace sur A d'un fermé de E.

Par passage au complémentaire, des propriétés de la topologie de A on déduit qu'une intersection quelconque et qu'une union finie de fermés de A sont encore des fermés de A. Si $u = (u_n)_{n \in \mathbb{N}}$ est une suite de points de A, on dit que cette suite converge dans A ssi elle est convergente dans E et si sa limite appartient à A. La caractérisation séquentielle des fermés de E donne le théorème suivant :

THÉORÈME 6-4.53 Une partie ϕ de A est un fermé relatif de A ssi toute suite de points de ϕ qui converge dans A a sa limite dans ϕ.

Démonstration : Exercice. ∎

EXERCICE 6-4.54 Montrer que tout fermé relatif de A est fermé de E ssi A est un fermé de E. Démontrer un résultat analogue pour les ouverts.

6-5 Suites de Cauchy, espaces complets

6-5.1 Suites de Cauchy

DÉFINITION 6-5.1 *Une suite $u = (u_n)_{n\in\mathbb{N}}$ d'un espace normé $(E, \| \|)$ est dite de Cauchy si et seulement si elle vérifie la propriété :*

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n, m \in \mathbb{N} \quad n \geqslant N_\varepsilon \text{ et } m \geqslant N_\varepsilon \Longrightarrow \|u_n - u_m\| \leqslant \varepsilon$$

Il est clair que si la norme est remplacée par une norme équivalente, les suites de Cauchy conservent cette propriété. Les propriétés suivantes sont conséquences immédiates de la définition :

- Toute suite convergente est de Cauchy.
- Toute suite extraite d'une suite de Cauchy est de Cauchy.
- Toute suite de Cauchy est bornée.

Dans $\mathbb{R}$ ou $\mathbb{C}$, les suites de Cauchy convergent. La démonstration de ce résultat a été basée sur l'existence d'une valeur d'adhérence. On a d'ailleurs la propriété générale :

THÉORÈME 6-5.2 Si une suite de Cauchy d'un espace normé possède une valeur d'adhérence, elle est convergente.

Démonstration : Exercice. ■

Il existe des espaces normés où les suites de Cauchy ne sont pas toutes convergentes. L'exercice suivant en donne un exemple :

EXERCICE 6-5.3 Dans l'espace $(C^0([0,1],\mathbb{R}), \| \|_1)$,montrer que la suite de fonctions $(f_n)_{n\in\mathbb{N}-\{0,1\}}$ définies par

$$f_n(x) = \begin{cases} 0 & \text{si } 0 \leqslant x \leqslant 1/2 - 1/n \\ n(x - \frac{1}{2}) + 1 & \text{si } 1/2 - 1/n \leqslant x \leqslant 1/2 \\ 1 & \text{si } 1/2 \leqslant x \leqslant 1 \end{cases}$$

est de Cauchy. Si elle convergeait (au sens de la norme) vers une fonction g, en étudiant $\|f_n - g\|_1$, montrer que g serait nulle sur $[0, \frac{1}{2}[$ et étudier de même g sur $[\frac{1}{2}, 1]$. Conclure.

EXERCICE 6-5.4 Si $u = (u_n)_{n\in\mathbb{N}}$ est une suite d'un espace normé, on pose

$$U_n = \{u_k,\ k \geqslant n\}$$

Montrer qu'une suite u est de Cauchy ssi les U_n sont bornés et si leur diamètre tend vers 0 pour $n \to +\infty$.

6-5.2 Espace de Banach

DÉFINITION 6-5.5 (Espace complet) *Un espace vectoriel normé est dit complet ssi toute suite de Cauchy de cet espace est convergente. Un espace vectoriel normé complet est aussi appelé* **espace de Banach**. *Cette propriété est évidemment conservée si on remplace la norme par une norme équivalente.*

THÉORÈME 6-5.6 Un produit d'espaces vectoriels normés muni de la norme du sup (notée précédemment N^∞) est complet ssi chacun des espaces est complet.

Démonstration : Il suffit de remarquer que la convergence d'une suite dans l'espace produit est la convergence coordonnée par coordonnée, et qu'une suite de l'espace produit est de Cauchy pour N^∞ ssi les suites coordonnées le sont dans leurs espaces respectifs. De plus une suite de Cauchy dans un des espaces de coordonnée peut être "remontée" en une suite de Cauchy du produit, en imposant par exemple à toutes les autres coordonnées d'être nulles. ■

COROLLAIRE 6-5.7 $\mathbb{K}^p$ muni d'une des trois normes "usuelles" est un espace de Banach.

EXERCICE 6-5.8 Montrer qu'un espace vectoriel normé est complet ssi toute suite décroissante pour l'inclusion de parties fermées bornées dont le diamètre tend vers 0 a une intersection réduite à un point. Cette propriété généralise d'une certaine manière la propriété de $\mathbb{R}$ connue sous le nom de théorème des segments emboîtés.

EXERCICE 6-5.9 Dans un espace de Banach, l'intersection d'une suite décroissante de boules fermées est une boule fermée (éventuellement de rayon nul, c'est-à-dire réduite à un point). On commencera par montrer que les suites des centres et des rayons convergent.

Terminons enfin par un exemple d'espace de Banach sur lequel nous reviendrons ultérieurement.

THÉORÈME 6-5.10 Si X est un ensemble non vide et $(E, \| \|)$ un evn, l'espace vectoriel normé $(\mathcal{B}(X,E), \| \|_\infty)$ est complet ssi $(E, \| \|)$ est un espace de Banach. En particulier l'espace des fonctions numériques bornées sur X est complet pour la norme de la convergence uniforme.

Démonstration : Si $(\mathcal{B}(X,E), \| \|_\infty)$ est complet, et si $u = (u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de $(E, \| \|)$, la suite de fonctions f_n constantes sur X de valeur u_n est évidemment de Cauchy dans $(\mathcal{B}(X,E), \| \|_\infty)$, puisque $\|f_n - f_m\|_\infty = \|u_n - u_m\|$. Elle converge donc vers une fonction $g \in \mathcal{B}(X,E)$ pour $\| \|_\infty$. En particulier, si $x_0 \in X$ est fixé, on a

$$\|u_n - g(x_0)\| = \|f_n(x_0) - g(x_0)\| \leqslant \|f_n - g\|_\infty$$

ce qui montre que la suite u converge dans $(E, \| \|)$ vers $g(x_0)$. E est donc complet. Réciproquement, si $(E, \| \|)$ est complet, et si $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de $(\mathcal{B}(X, E), \| \|_\infty)$, pour $x \in X$ on a

$$\|f_n(x) - f_m(x)\| \leqslant \|f_n - f_m\|_\infty$$

majoration qui montre que la suite $(f_n(x))_{n\in\mathbb{N}}$ est de Cauchy dans $(E, \| \|)$. Elle est donc convergente vers une limite qui dépend du choix de x. On note $l(x)$ cette limite. On définit ainsi une application $l : X \to E$. On termine la démonstration en montrant que l est dans $\mathcal{B}(X, E)$ et que la suite f_n converge vers g pour $\| \|_\infty$. La suite f_n étant de Cauchy dans $(\mathcal{B}(X, E), \| \|_\infty)$, elle est bornée. Il existe un réel M tel que pour tout n on ait $\|f_n\|_\infty \leqslant M$. On a alors

$$\forall\, x \in X \quad \forall\, n \in \mathbb{N} \quad \|f_n(x)\| \leqslant M$$

Pour x fixé et en faisant tendre n vers $+\infty$ on obtient $\|l(x)\| \leqslant M$, ce qui montre que $l \in \mathcal{B}(X, E)$. Enfin, si $\varepsilon > 0$ est donné, il existe un rang N_ε tel que

$$n \text{ et } m \geqslant N_\varepsilon \Longrightarrow \forall\, x \in X \quad \|f_n(x) - f_m(x)\| \leqslant \varepsilon$$

Si n est fixé arbitrairement $\geqslant N_\varepsilon$, en faisant tendre m vers $+\infty$ on obtient

$$n \geqslant N_\varepsilon \Longrightarrow \forall\, x \in X \quad \|l(x) - f_n(x)\| \leqslant \varepsilon$$

ce qui donne exactement la convergence de f_n vers l dans $(\mathcal{B}(X, E), \| \|_\infty)$. ■

6-5.3 Partie complète

DÉFINITION 6-5.11 (Partie complète) *Une partie A d'un espace normé E est dite complète si et seulement si toute suite de Cauchy de points de A est convergente* ***dans*** *A.*

THÉORÈME 6-5.12 Si A est une partie complète d'un evn E, A est fermée.

Démonstration : Si l est limite d'une suite convergente de points de A, cette suite est de Cauchy et converge donc dans A par hypothèse. On a donc $l \in A$ et A est fermé. ■

THÉORÈME 6-5.13 Dans un espace de Banach, les parties complètes sont exactement les parties fermées.

Démonstration : Si $A \subset E$ est fermée, une suite de Cauchy de points de A est convergente dans E complet, et sa limite appartient à A car A est fermée. ■

6-6 Exercices

EXERCICE 6-6.1 Soient (a_n) et (b_n) deux suites réelles convergentes. Etudier la suite (c_n) définie par :

$$c_n = \frac{1}{n+1}\sum_{i=0}^{n} a_i b_{n-i}$$

EXERCICE 6-6.2 Soit $f_n(x) = x^n + x^{n-1} + \cdots + x^2 + x - 1$ et a_n l'unique racine positive de l'équation $f_n(x) = 0$. Etudier la suite a_n.

EXERCICE 6-6.3 Soit $x \in \mathbb{R}^p$ $(p \geqslant 2)$. On définit une suite (x_n) d'éléments de $\mathbb{R}^p$ par $x_0 = x$ et

$$\forall i \in \{1, \cdots, p\} \quad x_{n+1,i} = \frac{1}{p-1}\sum_{j \neq i} x_{n,j}$$

où $x_{n,j}$ est la j-ième composante de x_n. Etudier la convergence de la suite (x_n).

EXERCICE 6-6.4 Soit u_n une suite bornée de réels telle que $\lim\limits_{n \to +\infty} (u_{n+1} - u_n) = 0$. Montrer que l'ensemble des valeurs d'adhérence de la suite u_n est un segment. Que se passe-t-il si la suite u_n n'est pas bornée ? Construire une suite réelle dont l'ensemble des valeurs d'adhérence est $\mathbb{R}$.
Soient u_n et v_n deux suites réelles telles que :

$$\lim_{n \to +\infty} u_n = \lim_{n \to +\infty} v_n = +\infty \text{ et } \lim_{n \to +\infty} (u_{n+1} - u_n) = 0$$

Montrer que $\{u_n - v_m \ : n, m \in \mathbb{N}\}$ est dense dans $\mathbb{R}$.

EXERCICE 6-6.5 On considère la suite définie par $u_0 > 0$ et $u_{n+1} = \dfrac{1}{n}\mathrm{Arctg}(nu_n)$

1. Montrer que la suite u_n converge.
2. Montrer que la suite nu_n est bornée, puis qu'elle ne possède qu'une seule valeur d'adhérence. Que peut on en déduire ?

EXERCICE 6-6.6 On note $l^2(\mathbb{N}) = \{(u_n)_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}} \mid \sum |u_n|^2 \text{ converge}\}$.

1. Montrer que $l^2(\mathbb{N})$ est un $\mathbb{C}$-ev et que $\|u\|_2 = \left(\sum |u_n|^2\right)^{\frac{1}{2}}$ est une norme sur $l^2(\mathbb{N})$.
2. Montrer que pour cette norme $l^2(\mathbb{N})$ est complet.

EXERCICE 6-6.7 Sur $\mathcal{C}^1([0,1], \mathbb{R})$ montrer que :

$$N_1(f) = |f(0)| + \sup_{[0,1]} |f + 2f'| \ \text{ et } \ N_2(f) = \sup_{[0,1]} |f| + \sup_{[0,1]} |f'|$$

sont deux normes équivalentes. Sont-elles équivalentes à la norme de la convergence uniforme ?

EXERCICE 6-6.8 Soit $(E, \| \ \|)$ un espace vectoriel normé complet. Montrer que si $(B_n)_{n \in \mathbb{N}}$ est une suite de boules fermées décroissante pour l'inclusion, alors $\bigcap\limits_n B_n$ est une boule fermée.

EXERCICE 6-6.9 On considère $\mathbb{R}^p$ muni de sa topologie usuelle. Soit F un voisinage fermé de 0, convexe, borné, invariant par symétrie par rapport à 0. Montrer qu'il existe sur $\mathbb{R}^p$ une norme unique dont F soit la boule unité fermée.

EXERCICE 6-6.10 On se propose de déterminer $S = \{(x,y) \in \mathbb{Z}^2 \mid x^2 - 2y^2 = 1\}$

1. Montrer que $G = \{z \in \mathbb{R} \mid z > 0 \text{ et } \exists (x,y) \in S \;\; z = x + y\sqrt{2}\}$ est un sous groupe multiplicatif de $\mathbb{R}^*$. Montrer que les éléments de G supérieurs à 1 vérifient $x \geqslant 0$ et $y \geqslant 0$.

2. En déduire G puis S.

EXERCICE 6-6.11 Soit ρ une semi-norme sur $M_n(\mathbb{C})$ vérifiant

$$\forall A, B \;\; \rho(AB) \leqslant \rho(A)\rho(B)$$

Montrer que ρ est identiquement nulle ou est une norme sur $M_n(\mathbb{C})$.

EXERCICE 6-6.12 Soit E l'espace des fonctions réelles bornées sur $[0,1]$ muni de $\|f\|_\infty = \sup\limits_{[0,1]} |f|$. Si A est une partie non vide de $[0,1]$ et $X = \{f \in E \mid f|_A = 0\}$, montrer que X est égal à sa frontière.

EXERCICE 6-6.13 Pour tout n on pose $u_n = \left(\dfrac{1+\sqrt{5}}{2}\right)^n$, et $a_n = d(u_n, \mathbb{Z})$. Quel est le comportement de la suite (a_n) ?

EXERCICE 6-6.14 Soit $E = \{(x,y) \in \mathbb{R}^2 \mid y^2(2-y^2) = x^2(x-1)(x-2)\}$. Montrer que E est borné pour $\| \; \|_\infty$.

EXERCICE 6-6.15 Soit (u_n) une suite réelle "sous-additive", c'est à dire que

$$\forall n,\, p,\; u_{n+p} \leqslant u_n + u_p$$

1. Soient m et n deux entiers strictement positifs tels que m divise n. Montrer que : $u_n/n \leqslant u_m/m$.

2. Soit $m \in \mathbb{N}^*$; on pose $\beta = \sup\{|u_r|,\; 0 \leqslant r \leqslant m\}$. Montrer que, pour tout $n \geqslant m$, on a
$$\frac{u_n}{n} \leqslant \frac{u_m}{m} + \frac{2\beta}{n}$$

3. En déduire que la suite (u_n/n) converge vers sa borne inférieure (si elle est minorée) ou diverge vers $-\infty$.

EXERCICE 6-6.16 Dans $\mathbb{R}^2$, on pose $\|x\| = 2|x_1| + 3|x_2|$. Vérifier que ceci définit une norme et trouver les meilleures constantes α et β, telles que :

$$\forall x,\; \alpha\|x\| \leqslant \|x\|_2 \leqslant \beta\|x\|$$

Même question dans $\mathbb{R}^4$ avec

$$\|x\| = \max(2|x_1| + 3|x_2|, 3|x_3| + 2|x_4|)$$

EXERCICE 6-6.17 Soient x_n et y_n deux suites réelles vérifiant $\forall n \geqslant 1 \;\; y_n = 2x_n + x_{n-1}$. Montrer que la suite x_n converge ssi la suite y_n est convergente.

EXERCICE 6-6.18 Etudier la suite définie par $u_1 > 0$ et $u_{n+1} = \sqrt{u_n + \sin \dfrac{1}{n}}$.

Chapitre 7

Continuité et Compacité

7-1 Limite et continuité

7-1.1 Limite et continuité en un point

Dans tout ce chapitre, E et F représentent deux espaces normés sur lesquels les normes sont représentées (s'il n'y a pas d'ambiguïté) par le même symbole $\| \ \|$, bien que ces deux normes soient en général distinctes. Cela ne doit pas créer de confusion, il suffit d'être attentif à l'espace dans lequel sont évaluées les normes.

DÉFINITION 7-1.1 *Soit $A \subset E$ et $f : A \to F$ une application, a étant un point de E adhérent à A. On dit que f admet une limite en a ssi*

$$\exists\, b \in F \quad \forall \varepsilon > 0 \quad \exists\, \alpha > 0 \quad \forall\, x \in A \quad \|x - a\| \leqslant \alpha \Longrightarrow \|f(x) - b\| \leqslant \varepsilon$$

Comme tout voisinage d'un point contient une boule ouverte centrée en ce point (qui est elle-même voisinage du point), cette définition équivaut à

$$\exists\, b \in F \quad \forall\, V \in \mathcal{V}(b) \quad \exists\, W \in \mathcal{V}(a) \quad f(W \cap A) \subset V$$

Il est à noter que l'hypothèse $a \in \overline{A}$ est indispensable pour donner un sens à cette définition, en assurant que si W est un voisinage de A, $W \cap A \neq \emptyset$. Il est clair également que si f admet une limite en a alors f est bornée sur un ensemble de la forme $W \cap A$ (il suffit de choisir le voisinage associé à $\varepsilon = 1$ sur lequel on a la majoration $\|f(x)\| \leqslant \|b\| + 1$).

THÉORÈME 7-1.2 (Unicité de la limite) Si f possède une limite en a, le vecteur b intervenant dans la définition précédente est unique, on le note

$$b = \lim_a f = \lim_{x \to a} f(x)$$

Démonstration : Si $b_1 \neq b_2$ répondent à cette définition, on peut en trouver des voisinages disjoints V_1 et V_2 (F est "séparé"), et leur associer des voisinages W_1 et W_2 de a avec

$$f(W_1 \cap A) \subset V_1 \text{ et } f(W_2 \cap A) \subset V_2$$

On arrive alors à une contradiction puisque

$$\emptyset \neq f(W_1 \cap W_2 \cap A) \subset V_1 \cap V_2 = \emptyset \quad \blacksquare$$

On remarque également que si $g : A \to F$ coïncide avec f sur un ensemble de la forme $W_0 \cap A$ où W_0 est un voisinage particulier de $a \in \overline{A}$, on a également $b = \lim_a g$, puisque pour $W \in \mathcal{V}(a)$ on a

$$f(W \cap W_0 \cap A) = g(W \cap W_0 \cap A)$$

On dit que l'existence d'une limite en a est une **propriété locale.**

La définition précédente amène à distinguer deux cas :

- Si $a \in A$ et si $f : A \to F$ possède une limite b en a, comme $f(a)$ doit appartenir à tout voisinage de b, on a forcément $f(a) = b$, et on dit que f est **continue en a**. La propriété de continuité peut se traduire par

$$\forall V \in \mathcal{V}(f(a)) \quad \exists W' \text{ voisinage de } a \text{ relatif à } A \quad f(W') \subset V$$

- Si $a \in \overline{A} - A$, et si $\lim_{x \to a} f(x) = b$, il existe un prolongement "naturel" $\widetilde{f}$ de f à $A \cup \{a\}$ obtenu en posant $\widetilde{f}(a) = b$. Cette application est le seul prolongement de f à $A \cup \{a\}$ qui soit continu en a (vérification facile), on dit que $\widetilde{f}$ est **le prolongement par continuité de f en a.**

7-1.2 Généralisation de la définition

7-1.2.1 Limite suivant une partie

DÉFINITION 7-1.3 *Soit $A \subset E$ et $f : A \to F$. Si $P \subset A$ et $a \in \overline{P}$, on dit que f possède une limite en a suivant P ssi la fonction $f|_P$ possède une limite en a et on note*

$$\lim_{\substack{x \to a \\ x \in P}} f(x) = \lim_{x \to a} f|_P(x)$$

cette limite. On a donc $\lim_{\substack{x \to a \\ x \in P}} f(x) = b$ ssi

$$\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall x \in P \quad \|x - a\| \leqslant \alpha \Longrightarrow \|f(x) - b\| \leqslant \varepsilon$$

Par exemple, si $f : \mathbb{R} \to \mathbb{R}$ est la fonction caractéristique de $\mathbb{R} - \mathbb{Q}$, on a $\lim_{\substack{x \to 0 \\ x \in \mathbb{Q}}} f(x) = 0$, alors que l'on a $\lim_{\substack{x \to 0 \\ x \in \mathbb{R} - \mathbb{Q}}} f(x) = 1$. La fonction f n'a évidemment pas de limite en 0.

7-1.2.2 Limite à l'infini

DÉFINITION 7-1.4 *Si A est une partie non bornée de E, on dit que $f : A \to F$ tend vers b à l'infini et on note $\lim_{\|x\| \to +\infty} f(x) = b$ si et seulement si :*

$$\forall \varepsilon > 0 \quad \exists\, \alpha > 0 \quad \forall\, x \in A \quad \|x\| \geqslant \alpha \Longrightarrow \|f(x) - b\| \leqslant \varepsilon$$

ce qui peut aussi se traduire en termes de voisinages par :

$$\forall\, V \in \mathcal{V}(b) \quad \exists\, W \in \mathcal{V}(\infty) \quad f(W \cap A) \subset V$$

Si $E = \mathbb{R}$, un voisinage de $+\infty$ (resp. $-\infty$) est une partie contenant une demi-droite de la forme $[M, +\infty[$ (resp. $]-\infty, M]$). On peut ainsi donner une définition correspondant à l'égalité $\lim_{x \to +\infty} f(x) = b$.

DÉFINITION 7-1.5 *Si $f : A \to \mathbb{R}$ et $a \in \overline{A}$, on dit que f tend vers $+\infty$ en a ssi*

$$\forall\, V \in \mathcal{V}(+\infty) \quad \exists\, W \in \mathcal{V}(a) \quad f(W \cap A) \subset V$$

ce qui traduit la propriété

$$\forall M > 0 \quad \exists\, \alpha > 0 \quad \forall\, x \in A \quad \|x - a\| \leqslant \alpha \Longrightarrow f(x) \geqslant M$$

On donnerait de manière similaire une définition pour $\lim_{\|x\| \to +\infty} f(x) = +\infty$ lorsque A n'est pas bornée.

Les définitions précédentes se généralisent à la notion de limite suivant une partie P de A.

7-1.3 Caractérisation séquentielle

THÉORÈME 7-1.6 Si $f : A \to F$ et $a \in \overline{A}$, f possède une limite en a égale à $b \in F$ ssi, pour tout suite $(a_n)_{n \in \mathbb{N}}$ de points de A convergeant vers a la suite des images $(f(a_n))_{n \in \mathbb{N}}$ converge vers b. En particulier f est continue en $a \in A$ si les suites images convergent toutes vers $f(a)$.

Démonstration : Si $b = \lim_a f$, si V est un voisinage arbitraire de b, il existe un voisinage W de a avec $f(W \cap A) \subset V$. Si $(a_n)_{n \in \mathbb{N}}$ est une suite de points de A convergeant vers a, à partir d'un certain rang tous les termes de la suite sont dans $W \cap A$, ce qui entraîne que tous les termes de la suite image sont, à partir du même rang, dans V. Ceci traduit bien

$$\lim_{n \to +\infty} f(a_n) = b$$

Réciproquement, si f ne tend pas vers b en a, on a

$$\exists\, \varepsilon > 0 \quad \forall\, \alpha > 0 \quad \exists\, x \in A \quad \|x - a\| \leqslant \alpha \text{ et } \|f(x) - b\| > \varepsilon$$

Pour ce choix de ε, et en prenant $\alpha = \dfrac{1}{n}$ avec $n \in \mathbb{N}^*$, on est assuré de l'existence d'un point $x_n \in A$, avec $\|x_n - a\| \leqslant \dfrac{1}{n}$ et $\|f(x_n) - b\| > \varepsilon$. La suite x_n converge bien vers a, mais la suite des images ne converge pas vers b. ■

On notera que, si pour toute suite de points de A qui converge vers a la suite des images est convergente dans F, alors la limite de cette suite image ne dépend pas de la suite choisie dans A. Il suffit pour s'en convaincre de prendre deux suites (a_n) et (a'_n) dans A qui tendent vers a et de les "mélanger" en une seule suite $b = (b_n)$ en posant $b_{2n} = a_n$ et $b_{2n+1} = a'_n$.

Les théorèmes sur les suites convergentes démontrés dans le chapitre précédent permettent d'obtenir les résultats suivants :

THÉORÈME 7-1.7 (Limite dans un espace produit) Si $f : A \to F_1 \times \cdots \times F_p$ est à valeurs dans l'espace normé produit des espaces $(F_i, \|\ \|)_{1 \leqslant i \leqslant p}$, la fonction $f = (f_1, \ldots, f_p)$ possède une limite $l = (l_1, \ldots, l_p)$ en $a \in \overline{A}$ ssi chacune des applications composantes f_i possède la limite l_i en a. En particulier, la fonction f est continue en $a \in A$ ssi toutes les applications composantes sont continues en a.

THÉORÈME 7-1.8 Les applications

$$\begin{array}{ll} E \times E \to E & (x, y) \mapsto x + y \\ \mathbb{K} \times E \to E & (\lambda, x) \mapsto \lambda x \\ E \to \mathbb{R}^+ & x \mapsto \|x\| \end{array}$$

sont continues en tout point de leur espace de départ (muni de sa structure naturelle d'espace produit pour les deux premières applications)

THÉORÈME 7-1.9 (Composition) Si $f : A \to F$ possède une limite égale à b en $a \in \overline{A}$, si $g : F \supset B \to G$ à valeurs dans l'evn G est telle que $f(A) \subset B$ et possède une limite égale à l en b, alors $g \circ f$ tend vers l en a.

> *Démonstration* : On remarque tout d'abord que si (a_n) une suite de points de A qui converge vers a, la suite $(f(a_n))$ est dans B et converge vers b qui est donc bien dans $\overline{B}$. Comme g a la limite l en b, la suite $(g(f(a_n))_{n \in \mathbb{N}}$ converge donc vers l. ■

On notera l'importance de l'hypothèse $f(A) \subset B$ dans le théorème précédent. Par exemple, si $g = \mathbf{1}_{\mathbb{R}^*}$ (fonction caractéristique de $\mathbb{R}^*$) et $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x \sin\left(\frac{1}{x}\right)$ pour $x \neq 0$ et $f(0) = 0$, on vérifie que

$$\lim_{\substack{x \to 0 \\ x \neq 0}} f(x) = 0 \text{ et } \lim_{\substack{x \to 0 \\ x \neq 0}} g(x) = 1$$

Par contre $\lim\limits_{\substack{x \to 0 \\ x \neq 0}} g \circ f(x)$ n'existe pas.

THÉORÈME 7-1.10 (Opérations algébriques) Si $f : A \to F$, $g : A \to F$ et $\alpha : A \to \mathbb{K}$ possèdent des limites en $a \in \overline{A}$, il en est de même de $f + \alpha g$ et on a :

$$\lim_a (f + \alpha g) = \lim_a f + \left(\lim_a \alpha\right)\left(\lim_a g\right)$$

Si f, g et α sont continues en $a \in A$ alors $f + \alpha g$ l'est également.

THÉORÈME 7-1.11 Les fonctions polynomiales $(\mathbb{K}^p, \| \|_\infty) \to \mathbb{K}$ sont continues en tout point de $\mathbb{K}^p$.

THÉORÈME 7-1.12 Si $f : A \to \mathbb{K}$ possède une limite non nulle en $a \in \overline{A}$, la fonction $\frac{1}{f}$ est définie au moins sur un voisinage de a relatif à A et

$$\lim_a \frac{1}{f} = \frac{1}{\lim_a f}$$

De même si f est continue en $a \in A$ avec $f(a) \neq 0$, la fonction $\frac{1}{f}$ est continue en a.

THÉORÈME 7-1.13 Les fonctions fractions rationnelles $\mathbb{K}^p \to \mathbb{K}$ sont continues en tout point de leur domaine de définition.

La continuité d'une application à valeurs dans un espace produit peut se lire, comme on l'a vu plus haut, sur chacune des applications composantes. Pour les fonctions définies sur un produit, la situation est plus complexe :

THÉORÈME 7-1.14 Si $E = E_1 \times E_2$ est un produit de deux espaces normés, $A_1 \subset E_1$ et $A_2 \subset E_2$, et $f : A_1 \times A_2 \to F$ est continue en $a = (a_1, a_2) \in A_1 \times A_2$, les applications partielles en a :

$$\begin{array}{lll} f(\bullet, a_2) : & A_1 \to F & x \mapsto f(x, a_2) \\ f(a_1, \bullet) : & A_2 \to F & y \mapsto f(a_1, y) \end{array}$$

sont continues respectivement en a_1 et a_2.

Démonstration : Il suffit d'appliquer la caractérisation séquentielle pour conclure immédiatement. ■

Il n'y a évidemment pas de réciproque à ce théorème, la connaissance des applications partielles en a ne donnant d'information sur f qu'en restriction à $A_1 \times \{a_2\} \cup \{a_1\} \times A_2$. Par exemple l'application

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad (x, y) \mapsto \frac{xy}{x^2 + y^2} \text{ avec } f(0,0) = 0$$

possède des applications partielles à l'origine continues en 0 (puisqu'identiquement nulles) alors que f n'est pas continue en $(0,0)$. En effet :

$$\lim_{\substack{x \to 0 \\ x \neq 0}} f(x, x) = \frac{1}{2} \neq f(0,0)$$

7-1.4 Utilisation d'espaces complets

Lorsque l'espace d'arrivée est un espace de Banach, le **critère de Cauchy** permet de démontrer l'existence d'une limite sans avoir à déterminer *a priori* la valeur de cette limite :

THÉORÈME 7-1.15 (Critère de Cauchy) Soit $f : A \to F$ espace de Banach et $a \in \overline{A} - A$. La fonction possède une limite en a ssi

$$\forall \varepsilon > 0 \quad \exists\, W \in \mathcal{V}(a) \quad \forall\, x, y \in W \cap A \quad \|f(x) - f(y)\| \leqslant \varepsilon \tag{1}$$

Ce résultat est encore valable si A n'est pas bornée pour étudier l'existence d'une limite éventuelle de f à l'infini.

Démonstration : L'inégalité triangulaire montre facilement que (1) est conséquence de l'existence d'une limite. Réciproquement, si la propriété (1) est vérifiée, il suffit de montrer que pour toute suite (a_n) de points de A qui converge vers a, la suite des images est convergente (*cf.* la remarque suivant le théorème **7-1**.6). Or si $\varepsilon > 0$ est donné, on dispose d'un voisinage W de a tel que le diamètre de $f(W \cap A)$ soit inférieur à ε. Si n et m sont suffisamment grands, a_n et a_m sont tous deux dans ce voisinage et vérifient donc $\|f(a_n) - f(a_m)\| \leqslant \varepsilon$, ce qui montre que la suite image est de Cauchy dans F espace complet, donc est convergente. ■

7-1.5 Continuité globale

DÉFINITION 7-1.16 *Une application $f : A \to F$ est dite (globalement) continue sur A ssi elle est continue en tout point de A. On notera $C^0(A, F)$ l'ensemble des applications continues de A dans F.*

Des théorèmes précédents découlent immédiatement les résultats :

- $C^0(A, F)$ est un sous-espace vectoriel de F^A. L'ensemble des fonctions numériques continues $C^0(A, \mathbb{K})$ est une $\mathbb{K}$-algèbre.
- Si $f : A \to F$ et $g : F \supset B \to G$ sont globalement continues avec $f(A) \subset B$, la composée $g \circ f$ est continue.
- Si $f : A \to F = \prod_{i=1}^{p} F_i$ a pour ensemble d'arrivée un produit d'espaces normés, f est continue ssi les applications composantes $p_i \circ f$ le sont, p_i étant la projection canonique $F \to F_i$, pour $1 \leqslant i \leqslant p$.
- Si $f : A \to F$ est continue, sa restriction à toute partie $P \subset A$ est continue.

THÉORÈME 7-1.17 (Continuité globale) Une application $f : A \to F$ est continue si et seulement si l'image réciproque de tout ouvert de F est un ouvert relatif de A. De manière équivalente, f est continue ssi l'image réciproque de tout fermé de F est un fermé de A.

Démonstration : Si f est continue et O est un ouvert de F, posons $\Omega = f^{-1}(O)$. Si Ω est vide, c'est un ouvert de A. Sinon, si $x \in \Omega$, O est un voisinage de $f(x)$. Par continuité de f en x, il existe donc un voisinage W de x dans E avec $f(W \cap A) \subset O$, soit $W \cap A \subset \Omega$. Ω est donc voisinage dans A de chacun de ses points, c'est un ouvert de A. Réciproquement, si l'image réciproque de tout ouvert de F est ouvert de A, montrons que f est continue en tout point $a \in A$. Si $\varepsilon > 0$, $f^{-1}(B(f(a), \varepsilon[)$ est un ouvert de A qui contient a, il contient donc un ensemble de la forme $A \cap B(a, \alpha[$, avec $\alpha > 0$. Ceci prouve la continuité de f en a. Si enfin X est un fermé de F, on a $f^{-1}(X) = A - f^{-1}(X^c)$, ce qui montre bien que la propriété de continuité globale se traduit également en termes d'images réciproques de fermés. ∎

REMARQUE 7-1.18 On n'a par contre aucune information en ce qui concerne l'image directe d'un ouvert ou d'un fermé de A par une application continue. Par exemple, la fonction $x \mapsto \sin x$ est continue sur $\mathbb{R}$, et $\sin(]0, 2\pi[) = [-1, 1]$ n'est pas ouvert de $\mathbb{R}$.

REMARQUE 7-1.19 Le théorème précédent est très souvent utilisé pour prouver que certaines parties d'un evn sont ouvertes ou fermées. Par exemple, dans $(\mathbb{K}^p, \| \|_\infty)$, si $P : \mathbb{K}^p \to \mathbb{K}$ est une application polynomiale

$$S_P = \{x = (x_1, \ldots, x_p) \mid P(x_1, \ldots, x_p) = 0\}$$

("hypersurface algébrique") est un fermé, puisque image réciproque de $\{0\}$ par la fonction continue P. C'est ainsi que l'on peut prouver que $GL_n(\mathbb{R})$ est un ouvert de $(M_n(\mathbb{R}), \| \|_\infty)$ sur lequel l'application $M \mapsto M^{-1}$ est continue.

EXERCICE 7-1.20 Montrer que $O_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) \mid {}^tMM = I_n\}$ est un fermé de $(M_n(\mathbb{R}), \| \|_\infty)$. Est-ce un ouvert de $(M_n(\mathbb{R}), \| \|_\infty)$?

EXERCICE 7-1.21 Montrer qu'une application $f : A \to F$ est continue ssi, pour toute partie $X \subset A$

$$f(\overline{X} \cap A) \subset \overline{f(X)}$$

On peut remarquer que dans cette caractérisation $\overline{X} \cap A$ est le plus petit fermé de A contenant X, c'est donc l'**adhérence de X dans A**.

7-1.6 Exemple d'utilisation de la continuité

Le fait de travailler avec des fonctions continues permet souvent, pour parler de manière vague, de "passer à la limite" (dans des égalités ou des inégalités). On a par exemple le résultat très important suivant :

THÉORÈME 7-1.22 Soit f et $g : A \to F$ deux applications continues coïncidant sur une partie $B \subset A$ dense dans A ($A \subset \overline{B}$). On a alors $f = g$.

Démonstration : Si a est un point de A, il existe une suite $(b_n)_{n\in\mathbb{N}}$ de points de B convergente vers a. Comme $\forall\ n \quad f(b_n) = g(b_n)$ et comme f et g sont continues en a on a

$$f(a) = \lim_{n\to+\infty} f(b_n) = \lim_{n\to+\infty} g(b_n) = g(a) \quad \blacksquare$$

L'exercice suivant est un bon exemple d'utilisation de ce genre d'argument (raisonnement "par densité").

EXERCICE 7-1.23 Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et si A et $B \in M_n(\mathbb{K})$, montrer que les matrices AB et BA ont même polynôme caractéristique.

Solution : Le résultat est clair si B est inversible puisqu'alors $AB = B^{-1}(BA)B$ est semblable à BA. Dans le cas général, on peut obtenir le résultat par un argument de continuité. Si on munit $M_n(\mathbb{K})$ de la norme $\| \|_\infty$, l'application

$$f : M_n(\mathbb{K}) \to \mathbb{K}^n \qquad B \mapsto \text{ le } n\text{-uplet des coefficients de } \chi_{AB}$$

(on a oublié le coefficient dominant $(-1)^n$), est continue puisque ses applications composantes sont évidemment polynomiales. Il en est de même de l'application g où χ_{AB} est remplacé par χ_{BA}. Or f et g coïncident sur $GL_n(\mathbb{K})$ dont on a montré (exemple **6-4**.39) qu'il est dense dans $M_n(\mathbb{K})$. On en déduit que g et f coïncident sur $M_n(\mathbb{K})$. Sur un corps quelconque, l'argument précédent ne pourrait être utilisé, mais on pourrait l'adapter en un raisonnement plus algébrique, en remarquant que les coefficients du polynôme $\chi_{A(B-XI_n)}$ sont des polynômes en X, qui coïncident avec ceux de $\chi_{(B-XI_n)A}$ hors du spectre de B. Si K est infini, un argument purement algébrique permet de conclure. Enfin, l'identité matricielle

$$\begin{pmatrix} A & -I_n \\ -XI_n & B \end{pmatrix} \begin{pmatrix} B & I_n \\ XI_n & 0 \end{pmatrix} = \begin{pmatrix} AB - XI_n & A \\ 0 & -XI_n \end{pmatrix}$$

donnerait une autre démonstration de ce résultat. $\blacksquare$

7-1.7 Notion d'homéomorphisme

DÉFINITION 7-1.24 (Homéomorphisme) *Deux espaces normés E et F étant donnés, une partie A de E est dite homéomorphe à une partie B de F ssi il existe une* **bijection** *$f : A \to B$* **bicontinue**, *c'est-à-dire que f et sa bijection réciproque f^{-1} sont continues. f est alors un homéomorphisme de A dans B.*

Il est à noter qu'une bijection continue n'a pas en général de réciproque continue. Par exemple, avec $E = \mathbb{R}$ et $F = \mathbb{C}$, l'application

$$f : [0, 1[\to \{z \in \mathbb{C} \mid |z| = 1\}$$

définie par $f(x) = e^{2i\pi x}$ est bijective, continue, mais sa réciproque n'est pas continue en 1.

EXERCICE 7-1.25 Montrer que la composée de deux homéomorphismes est un homéomorphisme. En déduire que la relation "être homéomorphe à" est une relation réflexive, symétrique et transitive.

EXERCICE 7-1.26 Si $f : A \to B$ est un homéomorphisme, montrer que l'application

$$\Omega \mapsto f(\Omega)$$

définit une bijection entre l'ensemble des ouverts relatifs de A et celui des ouverts relatifs de B. Un homéomorphisme entre A et B donne donc une correspondance bijective entre la topologie de A et celle de B.

7-2 Continuité uniforme

7-2.1 Continuité uniforme

DÉFINITION 7-2.1 (Continuité uniforme) *Si* $(E, \| \ \|)$ *et* $(F, \| \ \|)$ *sont deux espaces normés, une application* $f : E \supset A \to F$ *est dite uniformément continue ssi*

$$\forall\, \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall\, x, y \in A \quad \|x - y\| \leqslant \alpha \Longrightarrow \|f(x) - f(y)\| \leqslant \varepsilon \qquad (2)$$

On notera la place des quantificateurs : α dépend de ε, mais pas des points x et y, contrairement à la propriété de continuité en un point x quelconque de A. Il en résulte clairement qu'une application uniformément continue sur A est continue en tout point de A. Des inégalités sur les normes montrent que la propriété est conservée si l'on remplace la norme sur E ou F par une norme équivalente. La caractérisation suivante est souvent utilisée (par sa négation) pour prouver qu'une application n'est pas uniformément continue :

THÉORÈME 7-2.2 Une application $f : A \to F$ est uniformément continue ssi pour toutes les suites $(x_n, y_n)_{n\in\mathbb{N}}$ de $(A \times A)^{\mathbb{N}}$ telles que $\lim\limits_{n\to+\infty} \|x_n - y_n\| = 0$ on a $\lim\limits_{n\to+\infty} \|f(x_n) - f(y_n)\| = 0$

Démonstration : Si f est uniformément continue et $(x_n, y_n)_{n\in\mathbb{N}} \in (A \times A)^{\mathbb{N}}$ telles que l'on ait $\lim\limits_{n\to+\infty} \|x_n - y_n\| = 0$, si $\varepsilon > 0$ est fixé, la propriété (2) donne l'existence d'un réel $\alpha > 0$ tel que

$$\|x - y\| \leqslant \alpha \Longrightarrow \|f(x) - f(y)\| \leqslant \varepsilon$$

Si n est choisi assez grand on est assuré que $\|x_n - y_n\| \leqslant \alpha$ ce qui entraîne $\|f(x_n) - f(y_n)\| \leqslant \varepsilon$. On a donc bien

$$\lim_{n\to+\infty} \|f(x_n) - f(y_n)\| = 0$$

Réciproquement, si f n'est pas uniformément continue, en niant la propriété (2) on a

$$\exists\, \varepsilon > 0 \quad \forall \alpha > 0 \quad \exists\, x, y \in A \quad \|x - y\| \leqslant \alpha \text{ et } \|f(x) - f(y)\| > \varepsilon$$

Si on prend $\alpha = \frac{1}{n}$ avec $n \in \mathbb{N}^*$, on pourra donc trouver x_n et y_n dans A avec $\|x_n - y_n\| \leqslant \frac{1}{n}$ et $\|f(x_n) - f(y_n)\| > \varepsilon$. La suite $(\|f(x_n) - f(y_n)\|)_{n \in \mathbb{N}}$ ne tend pas vers 0. ■

Par exemple les suites $x_n = n$ et $y_n = n + \frac{1}{n}$ montrent que $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$, bien qu'elle soit continue en tout point de $\mathbb{R}$.

EXERCICE 7-2.3 Montrer qu'une composée d'applications uniformément continues est uniformément continue.

7-2.2 Applications lipschitziennes

On étudie dans ce paragraphe un cas particulièrement important d'applications uniformément continues.

DÉFINITION 7-2.4 (Applications lipchitziennes) *Soit $f : E \supset A \to F$ une application et $k \in \mathbb{R}^+$, on dit que f est k-lipschitzienne (ou lipschitzienne de rapport k) ssi*

$$\forall x, y \in A \quad \|f(x) - f(y)\| \leqslant k \|x - y\|$$

On remarquera que si f est k-lipschitzienne, elle est K-lipschitzienne pour tout $K > k$. D'où l'intérêt, pour une fonction dont on sait qu'elle est lipschitzienne, de trouver k le "meilleur possible", c'est-à-dire le plus petit possible. Il est clair (inégalités sur les normes) que remplacer les normes par des normes équivalentes ne change pas le caractère lipschitzien d'une application, mais la valeur du rapport est changée en général. Il est facile de voir qu'une composée d'applications lipschitziennes est lipschitzienne.

EXEMPLE 7-2.5 Les applications

$(E, \| \ \|) \to (\mathbb{R}, \| \ \|)$	$x \mapsto \|x\|$
$(E, \| \ \|) \to (\mathbb{R}, \| \ \|)$	$x \mapsto d(x, A)$ $(\emptyset \neq A \subset E)$
$(E, N^\infty) = \prod (E_i, N_i) \to (E_i, N_i)$	$x \mapsto p_i(x) = x_i$ ($i^{\text{ème}}$ projection)

sont 1-lipschitziennes.

THÉORÈME 7-2.6 Toute application lipschitzienne est uniformément continue.

Démonstration : Evident. ■

Il n'y a pas de réciproque à ce théorème. Par exemple, on verra que l'application $[0,1] \to \mathbb{R}$ définie par $x \mapsto \sqrt{x}$ est uniformément continue mais on voit facilement qu'elle n'est pas lipschitzienne. Pour les fonctions dérivables d'un intervalle de $\mathbb{R}$ dans $\mathbb{C}$, l'exercice suivant donne une caractérisation du caractère lipchitzien :

EXERCICE 7-2.7 Si I est un intervalle de $\mathbb{R}$, une application dérivable $f : I \to \mathbb{C}$ est k-lipschitzienne ssi

$$\forall x \in I \quad |f'(x)| \leqslant k$$

7-2.3 Théorème du point fixe

DÉFINITION 7-2.8 (Contraction) *Si f est une application de A dans F, on dit que f est une contraction ssi f est lipschitzienne de rapport $k < 1$. On dira aussi que f est k-contractante (ce qui suppose $k < 1$).*

Il faut remarquer que cette définition a un sens lorsque sont précisées les normes sur E et F. Si on remplace les normes par des normes équivalentes, la propriété n'est pas conservée en général.

THÉORÈME 7-2.9 (Point fixe) Si A est une partie complète d'un ev normé $(E, \| \|)$, et si $f : A \to A$ est une application k-contractante (pour la même norme $\| \|$ au départ et à l'arrivée), il existe dans A une unique solution l de l'équation

$$f(x) = x$$

Plus précisément, si x_0 est un point quelconque de A, la suite de points de A définie par la donnée de x_0 et la relation de récurrence

$$x_{n+1} = f(x_n)$$

converge vers l et on a les majorations :

$$\|x_n - l\| \leqslant \frac{k^n}{1-k} \|x_1 - x_0\| \quad \text{et} \quad \|x_n - l\| \leqslant \frac{k}{1-k} \|x_n - x_{n-1}\|$$

Démonstration : Il ne peut y avoir qu'un seul point fixe dans A, puisque si l et l' sont points fixes, on a

$$\|l - l'\| = \|f(l) - f(l')\| \leqslant k \|l - l'\|$$

donc $l = l'$, puisque $k < 1$. On montre ensuite l'existence en étudiant la convergence de la suite définie plus haut. Par récurrence, on montre que $\|x_{p+1} - x_p\| \leqslant k^p \|x_1 - x_0\|$. On a alors, pour $n < m$

$$\begin{aligned}\|x_m - x_n\| &\leqslant \|x_m - x_{m-1}\| + \cdots + \|x_{n+1} - x_n\| \\ &\leqslant (k^{m-1} + \cdots + k^n) \|x_1 - x_0\|\end{aligned}$$

et puisque $k < 1$ on a

$$\|x_m - x_n\| \leqslant \frac{k^n}{1-k} \|x_1 - x_0\| \tag{1}$$

inégalité qui montre que la suite (x_n) est de Cauchy dans l'espace complet A. Cette suite est donc convergente vers une limite $l \in A$, et comme f est continue en l (puisque f est lipschitzienne), on a

$$l = \lim_{n \to +\infty} x_{n+1} = \lim_{n \to +\infty} f(x_n) = f(l)$$

Par continuité de la norme, en faisant tendre m vers $+\infty$ dans l'inégalité (1), on obtient

$$\|x_n - l\| \leqslant \frac{k^n}{1-k} \|x_1 - x_0\|$$

et de même en considérant que la suite a été initialisée à l'instant $n-1$, on obtient

$$\|x_n - l\| \leqslant \frac{k}{1-k}\|x_n - x_{n-1}\|$$

Cette majoration justifie le test d'arrêt utilisé en général lorsqu'on programme le calcul d'une valeur approchée de l sur ordinateur. ■

Insistons sur les hypothèses de ce théorème : A doit être complète (donc si l'espace normé ambiant est un Banach, cela signifie A fermée) et f doit être contractante de A dans A. Ce théorème dont l'énoncé et la démonstration sont simples est à la base de très nombreux résultats en Analyse. Citons notamment le théorème d'existence et d'unicité locale pour les solutions d'une équation différentielle, le théorème des fonctions implicites, d'inversion locale. Dans ces trois exemples, la difficulté de la démonstration réside dans le choix de l'espace complet et d'une contraction "adaptée" au problème.

EXERCICE 7-2.10 A l'aide du théorème du point fixe, étudier la suite réelle définie par la donnée de u_0 et la relation de récurrence $u_{n+1} = \sin 2u_n$.

7-2.4 Exercice : prolongement d'une application uniformément continue

La notion de continuité uniforme a été introduite, historiquement, pour prouver l'intégrabilité au sens de Riemann d'une fonction continue sur un segment. Le théorème qui suit est une autre application très importante de la continuité uniforme :

EXERCICE 7-2.11 Si $f : A \to F$ est uniformément continue, montrer que l'image d'une suite de Cauchy de points de A est une suite de Cauchy dans F.

Ce résultat prend évidemment tout son intérêt lorsque F est complet. On a :

THÉORÈME 7-2.12 Si $A \subset (E, \| \ \|)$ et $f : A \to (F, \| \ \|)$ est uniformément continue à valeurs dans un espace de Banach, f se prolonge de manière unique en une fonction (uniformément) continue $\tilde{f}$ de $\overline{A}$ dans F. Ce prolongement est donc défini sur E dans le cas particulier où A est dense dans E.

Indication pour la démonstration : D'après (**7-1.6**) comme A est dense dans $\overline{A}$, deux fonctions continues sur $\overline{A}$ coïncidant sur A sont égales. Si le prolongement $\tilde{f}$ existe, il est unique. Pour le construire, on raisonne par analyse-synthèse. Si $a \in \overline{A}$, il existe une suite a_n de points de A convergeant vers a. Si $\tilde{f}$ existe, on a forcément $\tilde{f}(a) = \lim\limits_{n\to\infty} f(a_n)$. Montrer que cette suite converge dans F, que sa limite ne dépend pas de la suite choisie, et montrer enfin que le prolongement $\tilde{f}$ ainsi défini est uniformément continu sur $\overline{A}$. ■

Les applications de ce théorème sont multiples en analyse. Citons en particulier la construction de l'intégrale de Riemann que nous donnerons plus loin, la construction de l'intégrale de Lebesgue "abstraite", la transformation de Fourier etc...

7-3 Convergence uniforme et continuité

Rappelons des définitions et résultats vus précédemment. :

- Soit X un ensemble et E un espace normé. Sur l'espace vectoriel $\mathcal{B}(X,E)$ des applications bornées de X dans E, l'application $\| \ \|_\infty$ définie par

$$\|f\|_\infty = \sup\{\|f(x)\|\,,\ x \in X\}$$

est une norme, appelée **norme de la convergence uniforme** sur X. (théorème **6-2**.16)

- Si $(E, \| \ \|)$ est complet, il en est de même de $(\mathcal{B}(X,E), \| \ \|_\infty)$. (théorème **6-5**.10).

Dans cette section, nous allons préciser quelques propriétés fondamentales de la convergence uniforme, en relation avec la notion de continuité.

7-3.1 Convergence simple, convergence uniforme

DÉFINITION 7-3.1 *Soit X un ensemble, $(E, \| \ \|)$ un espace normé et $(f_n)_{n\in\mathbb{N}}$ une suite d'applications $X \to E$. On dit que la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur X vers une fonction $f : X \to E$ si et seulement si*

$$\forall x \in X \quad \lim_{n\to+\infty} f_n(x) = f(x)$$

Lorsque la convergence a lieu pour tout x d'une partie $A \subset X$, on parlera de convergence simple sur A.

L'étude de la convergence simple d'une suite de fonctions $X \to E$ est donc l'étude de la convergence de suites de vecteurs de E dépendant d'un paramètre $x \in X$. On notera cette convergence par

$$f_n \overset{\text{C.V.S}}{\longrightarrow} f$$

ce qui peut se traduire par

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \|f_n(x) - f(x)\| \leqslant \varepsilon$$

On notera bien la place des quantificateurs : N **dépend à la fois de x et ε**.

DÉFINITION 7-3.2 *La suite f_n convergence uniformément vers f sur X (en abrégé $f_n \overset{C.V.U}{\longrightarrow} f$) si et seulement si*

$$\lim_{n\to+\infty} \sup_{x\in X} \|f_n(x) - f(x)\| = 0$$

soit

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in X \quad \forall n \geqslant N \quad \|f_n(x) - f(x)\| \leqslant \varepsilon$$

Ici, l'ordre des quantificateurs montre que N ne dépend que de ε : à partir du rang N, la majoration $\|f_n(x) - f(x)\| \leqslant \varepsilon$ est valable pour tout $x \in X$.

Attention ! La norme $\| \|_\infty$ est définie sur l'espace des fonctions bornées de X dans E. La convergence uniforme de f_n vers f suppose évidemment que les différences $f_n - f$ soient bornées (à partir d'un certain rang) et tendent vers 0 dans $(\mathcal{B}(X,E), \| \|_\infty)$. Les fonctions f_n et f ne sont cependant pas nécessairement bornées sur X.

PROPOSITION 7-3.3 Si $f_n \overset{\text{C.V.U}}{\longrightarrow} f$ sur un ensemble X, alors $f_n \overset{\text{C.V.S}}{\longrightarrow} f$ sur X. La réciproque est fausse.

Par exemple, la suite de fonctions $f_n : [0,1[\to \mathbb{R}$ définie par $f_n(x) = x^n$ converge simplement vers la fonction nulle. La convergence n'est pas uniforme puisque

$$\forall n \in \mathbb{N} \quad \sup_{[0,1[} |f_n(x)| = 1$$

Pour avoir $|x^n| \leqslant \varepsilon$ avec $x \in]0,1[$, il faut $n \geqslant \dfrac{\ln \varepsilon}{\ln x}$. Pour obtenir un rang N qui convienne pour tous les x, il faudrait prendre

$$N \geqslant \frac{\ln \varepsilon}{\ln x} \text{ pour tout } x \in]0,1[$$

ce qui n'est pas possible, puisque $\displaystyle\lim_{x\to 1^-} \frac{\ln \varepsilon}{\ln x} = +\infty$.

REMARQUE 7-3.4 On dira que la suite f_n converge uniformément vers f sur une partie $A \subset X$ si et seulement si

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in A \quad \forall n \geqslant N \quad \|f_n(x) - f(x)\| \leqslant \varepsilon \qquad (1)$$

ce qui revient en fait à n'étudier que les restrictions des f_n et de f à A. Il est alors clair que la convergence uniforme sur des parties $A_1, A_2, \ldots, A_p$ de X entraîne la convergence uniforme sur la réunion $\bigcup_{i=1}^{p} A_i$. En effet, si $\varepsilon > 0$ est fixé, la propriété (1) appliquée à la partie A_i donne un entier N_i tel que

$$\forall n \geqslant N_i \quad \forall x \in A_i \quad \|f_n(x) - f(x)\| \leqslant \varepsilon$$

Pour $N = \max(N_1, \ldots, N_p)$, on aura bien

$$\forall n \geqslant N \quad \forall x \in \bigcup_{i=1}^{p} A_i \quad \|f_n(x) - f(x)\| \leqslant \varepsilon$$

ce qui prouve qu'il y a bien convergence uniforme sur $\bigcup_{i=1}^{p} A_i$. **Le raisonnement précédent n'est pas valable dans le cas d'une réunion d'une famille infinie de parties de X.** Par exemple, en reprenant les fonctions $f_n : [0,1[\to \mathbb{R}$

définies par $f_n(x) = x^n$, il y a bien convergence uniforme sur tout intervalle $[0, a] \subset [0, 1[$, puisque

$$\sup_{[0,a]} |f_n(x)| = a^n \underset{n \to +\infty}{\to} 0$$

Il n'y a pas convergence uniforme sur $\bigcup_{p \in \mathbb{N}^*} \left[0, 1 - \frac{1}{p}\right] = [0, 1[$

EXERCICE 7-3.5 Etudier la convergence simple et uniforme de la suite de fonctions $f_n : [0, +\infty[\to \mathbb{R}$ définies par

$$f_n(x) = n^\alpha x e^{-nx}$$

Discuter la convergence uniforme sur $\mathbb{R}^+$ en fonction du paramètre α. Montrer qu'il y a toujours convergence uniforme sur $[a, +\infty[$ pour tout $a > 0$.

7-3.2 Continuité d'une limite uniforme

On suppose à présent que X est une partie d'un espace normé. Une limite simple de fonctions continues $f_n : X \to E$ n'est pas forcément continue, comme le montre l'exemple $X = [0, 1]$ et $f_n(x) = x^n$. La limite simple est $f = 1_{\{1\}}$, fonction caractéristique du singleton $\{1\}$, discontinue en 1, bien que toutes les fonctions f_n soient continues en ce point.

La situation est autre en cas de convergence uniforme :

THÉORÈME 7-3.6 Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions $X \to E$ qui converge uniformément sur X vers une fonction f. Si toutes les f_n sont continues en un point $a \in X$, alors f est continue en a.

Démonstration : Si $\varepsilon > 0$ est donné, on peut trouver un entier N avec

$$\|f - f_N\|_\infty \leqslant \frac{\varepsilon}{3}$$

La fonction f_N étant continue en a, on peut trouver $\alpha > 0$ tel que

$$\forall x \in X \quad \|x - a\| \leqslant \alpha \Rightarrow \|f_N(x) - f_N(a)\| \leqslant \frac{\varepsilon}{3}$$

Pour $x \in X$ vérifiant $\|x - a\| \leqslant \alpha$, on a alors

$$\begin{aligned} &\|f(x) - f(a)\| \\ &\leqslant \|f(x) - f_N(x)\| + \|f_N(x) - f_N(a)\| + \|f_N(a) - f(a)\| \\ &\qquad \leqslant 2\|f - f_N\|_\infty + \|f_N(x) - f_N(a)\| \leqslant \varepsilon \end{aligned}$$

ce qui prouve bien la continuité de f en a. ∎

On notera bien, dans la démonstration précédente, l'importance de l'hypothèse de convergence uniforme : la différence $f(x) - f_N(x)$ peut être majorée pour un bon choix de N indépendamment de x. On peut ensuite imposer à x d'être dans un voisinage bien choisi de a, en utilisant la continuité de f_N.

REMARQUE 7-3.7 Comme la continuité est une propriété locale, la convergence uniforme sur un voisinage de a est suffisante pour assurer la conclusion du théorème.

COROLLAIRE 7-3.8 Une limite uniforme sur X d'une suite de fonctions continues sur X est continue sur X.

COROLLAIRE 7-3.9 L'espace $\mathcal{C}_b^0(X,E)$ des fonctions continues bornées de X dans E est fermé dans $(\mathcal{B}(X,E), \|\ \|_\infty)$. Il est donc complet (car fermé dans un complet) lorsque E est complet.

7-3.3 Théorème d'interversion des limites

L'interversion de passages à la limite ne peut être faite sans précautions. Par exemple, pour $f_n : [0,1[\to \mathbf{R}$ définie par $f_n(x) = x^n$

$$0 = \lim_{\substack{x\to 1\\ <}} \lim_{n\to+\infty} f_n(x) \neq \lim_{n\to+\infty} \lim_{\substack{x\to 1\\ <}} f_n(x) = 1$$

C'est souvent un argument de convergence uniforme qui permettra d'intervertir des passages à la limite.

THÉORÈME 7-3.10 Soit E un espace normé complet, A une partie d'un autre espace normé et $f_n : A \to E$ une suite de fonctions convergeant uniformément sur A vers une fonction f. Soit a un point de $\overline{A} - A$, adhérent à A. On suppose que chacune des fonctions f_n possède une limite $l_n \in E$ en a. Alors

- **La suite (l_n) est convergente dans E.**
- **La fonction f possède une limite en a et on a**

$$\lim_{x\to a} f(x) = \lim_{x\to a} \lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} l_n = \lim_{n\to+\infty} \lim_{x\to a} f_n(x)$$

Démonstration : Comme la suite $(f_n - f)$ est convergente dans $(\mathcal{B}(A,E), \|\ \|_\infty)$, elle est de Cauchy. Si $\varepsilon > 0$ est choisi arbitrairement, on peut trouver un rang N_ε tel que

$$\forall n \geqslant m \geqslant N_\varepsilon \quad \forall x \in A \quad \|f_n(x) - f_m(x)\| \leqslant \varepsilon$$

En faisant tendre x vers a, on obtient

$$\forall n \geqslant m \geqslant N_\varepsilon \quad \|l_n - l_m\| \leqslant \varepsilon$$

ce qui montre que la suite (l_n) est de Cauchy, donc converge vers une limite l dans l'espace complet $(E, \|\ \|)$. Définissons alors $X = A \cup \{a\}$ et prolongeons par continuité les fonctions f_n à X en posant $f_n(a) = l_n$. Prolongeons également f à a en posant $f(a) = l$. Par hypothèse, la suite $f_n \overset{\text{C.V.U}}{\longrightarrow} f$ sur A. Il y a aussi convergence (uniforme!) sur

$\{a\}$, donc finalement convergence uniforme sur X. Toutes les f_n étant continues en a, il en est de même de f, ce qui prouve que

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = l$$

et démontre le théorème d'interversion des limites. ■

REMARQUE 7-3.11 Il suffit en fait d'avoir convergence uniforme sur la trace sur A d'un voisinage de a.

REMARQUE 7-3.12 Dans la pratique, on n'utilise qu'une partie de la conclusion du théorème : on prouve en général d'abord l'existence des limites

$$\lim_{x \to a} \lim_{n \to +\infty} f_n(x) \quad \text{et} \quad \lim_{n \to +\infty} \lim_{x \to a} f_n(x)$$

et on utilise le théorème pour prouver l'égalité de ces deux termes.

REMARQUE 7-3.13 Si A est une partie non bornée de l'espace normé de départ, une légère modification des démonstrations des deux théorèmes précédents montrerait qu'on peut appliquer aussi le théorème d'interversion pour des limites à l'infini.

7-4 Applications linéaires continues

Dans tout ce paragraphe, $(E, \| \|)$ et $(F, \| \|)$ sont deux $\mathbb{K}$-espaces normés.

7-4.1 Caractérisation de la continuité d'une application linéaire

THÉORÈME 7-4.1 Si $f \in \mathcal{L}(E, F)$ est une application linéaire, les propriétés suivantes sont équivalentes

1. **f est uniformément continue sur E.**
2. **f est continue en 0_E.**
3. **f est bornée sur $B(0_E, 1]$.**
4. **f est bornée sur la sphère unité.**
5. $\exists\, k \in \mathbb{R}^+ \quad \forall x \in E \quad \|f(x)\| \leqslant k\,\|x\|$
6. $\exists\, k \in \mathbb{R}^+$ **avec f k-lipschitzienne.**

Démonstration : 1) $\Rightarrow$ 2), 3) $\Rightarrow$ 4), 5) $\Rightarrow$ 6) et 6) $\Rightarrow$ 1) sont évidents. Si f est continue en 0_E, il existe $\alpha > 0$ tel que

$$\|x\| \leqslant \alpha \Rightarrow \|f(x)\| \leqslant 1$$

Si $\|x\| \leqslant 1$, on aura donc $\|f(\alpha x)\| \leqslant 1$ et donc $\|f(x)\| \leqslant \dfrac{1}{\alpha}$. On a donc bien 2) $\Rightarrow$ 3). Enfin, si on suppose f bornée sur la sphère unité, et si k est la borne supérieure de $\|f(x)\|$ pour x parcourant la sphère unité, pour $x \neq 0_E$ on a $\left\|f\left(\dfrac{x}{\|x\|}\right)\right\| \leqslant k$ et donc $\|f(x)\| \leqslant k\|x\|$, inégalité qui est encore valable lorsque x est nul. On a prouvé 4) $\Rightarrow$ 5). ■

EXERCICE 7-4.2 Etudier la continuité des applications linéaires :

1. $f \mapsto \int_0^1 f(t)dt$ sur $(\mathcal{C}^0([0,1],\mathbb{R}), \| \|_\infty)$
2. $f \mapsto f'$ sur $(\mathcal{C}^1([0,1],\mathbb{R}), \| \|_\infty)$ à valeurs dans $(\mathcal{C}^0([0,1],\mathbb{R}), \| \|_\infty)$.
3. $P \mapsto P(2)$ sur $\mathbb{R}[X]$ muni de la norme $\|P\| = \sup_{t\in[0,1]} |P(t)|$.

On retrouve le résultat du chapitre précédent relatif à la comparaison des normes :

COROLLAIRE 7-4.3 Si E est un $\mathbb{K}$-ev et si N et N' sont deux normes, l'application $id_E : (E,N) \to (E,N')$ est continue ssi il existe $k \geqslant 0$ tel que $N' \leqslant kN$. Si les normes sont équivalentes, cette application est un homéomorphisme, ce qui traduit que les espaces (E,N) et (E,N') ont les mêmes ouverts.

On remarque que la continuité de $id_E : (E,N) \to (E,N')$ traduit exactement le fait que toute suite tendant vers 0_E dans (E,N) converge vers 0_E pour N'.

7-4.2 Espace $\mathcal{L}_c(E,F)$

7-4.2.1 Définition

DÉFINITION 7-4.4 *Si $(E,\| \|)$ et $(F,\| \|)$ sont deux $\mathbb{K}$-espaces normés, on note $\mathcal{L}_c(E,F)$ l'espace des applications linéaires* **continues** *de E dans F.*

THÉORÈME 7-4.5 $\mathcal{L}_c(E,F)$ est un sous-espace vectoriel de $\mathcal{L}(E,F)$.

Démonstration : L'application identiquement nulle est évidemment continue. Les théorèmes sur les combinaisons linéaires de fonctions continues donnent immédiatement le résultat. ■

DÉFINITION 7-4.6 *En particulier, on appelle* **dual topologique** *et on note en général E' l'ensemble des formes linéaires continues sur l'evn E ($\mathbb{K}$ est évidemment muni de sa topologie usuelle).*

Par exemple, toute forme linéaire sur $(\mathbb{K}^p, \| \|_\infty)$ est continue, puisqu'elle s'exprime comme polynôme homogène de degré 1 des coordonnées dans la base canonique. Par contre, sur un evn de dimension infinie, il existe des formes linéaires non continues :

EXEMPLE 7-4.7 Sur $\mathbb{R}[X]$, on vérifie aisément que

$$\|a_0 + a_1 X + \cdots + a_p X^p\| = \max |a_i|$$

définit une norme. Construire une forme linéaire non continue sur cet espace en la définissant sur les vecteurs de la base canonique. $P \mapsto P\left(\frac{1}{2}\right)$ et $P \mapsto \int_0^1 P(t)dt$ sont-elles continues ?

EXERCICE 7-4.8 Si φ est une forme linéaire sur E, montrer que φ est continue ssi $\ker \varphi$ est fermé.

Indication : $\ker \varphi = \varphi^{-1}(0)$. Une des implications en découle aisément. Si $H = \ker \varphi$ est fermé, et si $E = \ker \varphi \oplus \operatorname{vect} a$, montrer que, pour $x \in H$,

$$\|x + \lambda a\| \geqslant |\lambda| \, d(a, H)$$

En déduire la continuité de φ. ■

7-4.2.2 Norme d'une application linéaire continue

THÉORÈME 7-4.9 (et définition) Si $f \in \mathcal{L}_c(E, F)$, on définit la norme de l'application linéaire continue f par

$$\|f\| = \sup_{\substack{x \in E \\ x \neq 0_E}} \frac{\|f(x)\|}{\|x\|} = \sup_{\substack{x \in E \\ \|x\|=1}} \|f(x)\| = \sup_{\|x\| \leqslant 1} \|f(x)\|$$

c'est donc la plus petite constante $k \geqslant 0$ telle que

$$\forall x \in E \quad \|f(x)\| \leqslant k \|x\|$$

$(\mathcal{L}_c(E, F), \| \ \|)$ **est un espace normé.**

Démonstration : On prendra garde que, dans la définition précédente, il y a trois symboles $\| \ \|$, qui représentent des normes sur des espaces distincts en général. L'existence des bornes supérieures et leur égalité est conséquence de la caractérisation de la continuité d'une application linéaire et de l'homogénéité de la norme. Prouver qu'on définit ainsi une norme sur $\mathcal{L}_c(E, F)$ est élémentaire. Par exemple, pour l'inégalité triangulaire :

$$\begin{aligned} \forall x \in E \quad \|(f+g)(x)\| &\leqslant \|f(x)\| + \|g(x)\| \\ &\leqslant \|f\| \, \|x\| + \|g\| \, \|x\| = (\|f\| + \|g\|) \, \|x\| \end{aligned}$$

donne immédiatement $\|f + g\| \leqslant \|f\| + \|g\|$. ■

REMARQUE 7-4.10 Si on remplace la norme $\| \ \|$ sur E par une norme équivalente N, et celle sur F par une norme N' équivalente, on a des inégalités

$$\alpha N \leqslant \| \ \| \leqslant \beta N \text{ et } \alpha' N' \leqslant \| \ \| \leqslant \beta' N'$$

(attention à la signification des différents symboles), ce qui donne facilement, pour un vecteur non nul de E et $f \in \mathcal{L}(E,F)$

$$\frac{\alpha'}{\beta}\frac{N'(f(x))}{N(x)} \leqslant \frac{\|f(x)\|}{\|x\|} \leqslant \frac{\beta'}{\alpha}\frac{N'(f(x))}{N(x)} \tag{1}$$

L'espace $\mathcal{L}_c(E,F)$ n'a pas changé, puisque la continuité est une notion topologique. En travaillant avec les espaces (E,N) et (E',N'), on définirait une norme N'' sur $\mathcal{L}_c(E,F)$ par

$$N''(f) = \sup_{x \neq 0_E} \frac{N'(f(x))}{N(x)}$$

N'' est alors évidemment équivalente à la norme précédemment définie puisque (1) donne immédiatement

$$\frac{\alpha'}{\beta}N''(f) \leqslant \|f\| \leqslant \frac{\beta'}{\alpha}N''(f)$$

Lorsqu'on parle de norme d'une application linéaire continue, il est indispensable (s'il y a des risques d'ambiguïté) de préciser quelles normes ont été choisies au départ et à l'arrivée. On dit donc que la norme sur $\mathcal{L}_c(E,F)$ définie au théorème (**7-4**.9) est **la norme subordonnée aux normes** $\| \; \|$ **et** $\| \; \|$ (respectivement sur E et F).

EXERCICE 7-4.11 On considère $E = \mathbb{R}[X]$ muni de

$$\|a_0 + a_1X + \cdots + a_pX^p\| = \max |a_i|$$

et la forme linéaire $P \mapsto P(c)$ où c est un réel fixé. A quelle condition cette forme est-elle continue et déterminer alors sa norme, ($\mathbb{R}$ étant évidemment muni de $| \; |$). Réponse : $\dfrac{1}{1-|c|}$ si $|c| < 1$.

THÉORÈME 7-4.12 (Composition) Si $f \in \mathcal{L}_c(E,F)$ **et** $g \in \mathcal{L}_c(F,G)$, **alors** $g \circ f \in \mathcal{L}_c(E,G)$ **et**

$$\|g \circ f\| \leqslant \|g\| \, \|f\|$$

Démonstration : La composée de deux applications linéaires continues est évidemment linéaire et continue. De plus, en étant cohérent sur la signification des normes :

$$\forall x \in E \quad \|g \circ f(x)\| \leqslant \|g\| \, \|f(x)\| \leqslant \|g\| \, \|f\| \, \|x\|$$

ce qui donne immédiatement la majoration souhaitée. ∎

7-4.2.3 Convergence en norme

DÉFINITION 7-4.13 *Si $\mathcal{L}_c(E,F)$ est muni de la norme des applications linéaires continues subordonnée à des normes $\| \ \|$ sur E et F, on dit qu'une suite d'applications linéaires continues $f_n \in \mathcal{L}_c(E,F)$ converge en norme vers une application $f \in \mathcal{L}_c(E,F)$ (on dit aussi converge* ***fortement*** *vers f) ssi*

$$\lim_{n\to+\infty} \|f_n - f\| = 0$$

Il ne s'agit pas d'autre chose que de la convergence dans l'espace normé $(\mathcal{L}_c(E,F), \| \ \|)$.

THÉORÈME 7-4.14 Si $f_n \in \mathcal{L}_c(E,F)$ converge en norme vers $f \in \mathcal{L}_c(E,F)$, alors

$$\forall\, x \in E \quad \lim_{n\to+\infty} f_n(x) = f(x)$$

(on dit que la convergence forte entraîne la convergence simple). La réciproque est fausse en général.

Démonstration : Il suffit de remarquer que

$$\|f_n(x) - f(x)\| \leqslant \|f_n - f\|\,\|x\| \quad \blacksquare$$

Le contre-exemple suivant montre que la réciproque de ce théorème est fausse (bien qu'elle soit vraie pour les espaces normés de dimension finie comme on le verra plus loin en exercice). On note $E = c_0(\mathbb{N})$ le sous-espace de $l^\infty(\mathbb{R})$ formé des suites réelles convergentes vers 0. On le munit de la norme induite par $\| \ \|_\infty$, qu'on notera de la même manière. On définit une forme linéaire φ_n sur E par

$$\varphi_n(u) = u_n \text{ pour } u = (u_n)_{n\in\mathbb{N}}$$

On vérifie aisément que φ_n est continue sur E et vérifie $\|\varphi_n\| = 1$. La suite φ_n converge simplement sur E vers la forme linéaire nulle, mais ne converge évidemment pas en norme vers cette limite.

Lorsque l'espace d'arrivée est complet, il en est de même de l'espace des applications linéaires continues :

THÉORÈME 7-4.15 Lorsque l'espace F est complet, l'espace $(\mathcal{L}_c(E,F), \| \ \|)$ est un espace de Banach. En particulier, le dual topologique d'un espace normé est toujours complet.

Démonstration : Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de l'espace normé $(\mathcal{L}_c(E,F), \| \ \|)$. Pour montrer que cette suite est convergente, on commence par montrer, conformément à ce qui précède, que la suite est simplement convergente vers une application linéaire f de E vers F. On montrera ensuite que f est continue et que f_n converge en norme vers f. On a

$$\forall\, \varepsilon > 0 \quad \exists\, N_\varepsilon \in \mathbb{N} \quad \forall\, n, m \geqslant N_\varepsilon \quad \|f_n - f_m\| \leqslant \varepsilon$$

On a en particulier, pour $x \in E$ quelconque et $\varepsilon > 0$ fixé

$$\forall\, n, m \geqslant N_\varepsilon \quad \|f_n(x) - f_m(x)\| \leqslant \|f_n - f_m\| \, \|x\| \leqslant \varepsilon \, \|x\| \qquad (2)$$

ce qui montre que la suite $f_n(x)$ est de Cauchy, donc convergente dans l'espace complet $(F, \| \, \|)$. On note sa limite $f(x)$, puisqu'elle dépend de $x \in E$. On définit ainsi une correspondance de E vers F, clairement linéaire (il suffit en effet de passer à la limite dans les équations traduisant la linéarité des applications f_n). Si n est fixé $\geqslant N_\varepsilon$, la propriété (2) donne, en faisant tendre m vers $+\infty$ et par continuité de la norme,

$$\|f_n(x) - f(x)\| \leqslant \varepsilon \, \|x\|$$

ce qui montre que $f_n - f$ est continue, donc aussi $f = (f - f_n) + f_n$, et que $\|f_n - f\| \leqslant \varepsilon$ pour tout $n \geqslant N_\varepsilon$. Ceci montre bien que f est limite forte de la suite f_n. ■

REMARQUE 7-4.16 La convergence forte dans $\mathcal{L}_c(E, F)$ peut se traduire par

$$\lim_{n \to +\infty} \sup\{\|f_n(x) - f(x)\| \,, \quad x \in B(0_E, 1]\} = 0$$

c'est donc la **convergence uniforme sur la boule unité,** ce qui entraîne par homothétie et translation (les fonctions sont linéaires) la convergence uniforme sur toute partie bornée de E.

EXERCICE 7-4.17 (qu'il vaut mieux traiter après avoir vu les séries). Décrire le dual topologique de $\mathbb{R}[X]$ muni de la norme $\|a_0 + a_1 X + \cdots + a_p X^p\| = \max |a_i|$, qu'on notera dans la suite $\| \, \|_\infty$.

Indication : Si φ est une forme linéaire, elle est définie par l'image de la base canonique de $\mathbb{R}[X]$, c'est-à-dire par une suite de réels

$$(\alpha_n = \varphi(X^n))_{n \in \mathbb{N}}$$

Montrer que φ est continue ssi les sommes

$$S_n = \sum_{i=0}^{n} |\alpha_i|$$

sont majorées (indépendamment de n), ce qui veut dire que la série $\sum_{n=0}^{+\infty} \alpha_n$ est absolument convergente. Montrer que l'on a alors $\|\varphi\| = \sum_{n=0}^{+\infty} |\alpha_n|$ et que réciproquement une série de réels absolument convergente permet de définir une forme linéaire continue sur $\mathbb{R}[X]$. Cet exercice montre en particulier que l'espace $l^1(\mathbb{N})$ des séries absolument convergentes est complet pour la norme

$$\left\|(\alpha_n)_{n \in \mathbb{N}}\right\|_1 = \sum_{n=0}^{+\infty} |\alpha_n| \quad ■$$

7-4.3 Applications bilinéaires continues

7-4.3.1 Caractérisation de la continuité

THÉORÈME 7-4.18 **Si $(E, \| \|)$, $(F, \| \|)$ et $(G, \| \|)$, une application bilinéaire $B : E \times F \to G$ étant donnée et $E \times F$ étant muni d'une des normes usuelles définissant la topologie produit, les propriétés suivantes sont équivalentes :**

1. **B est continue sur $E \times F$.**
2. **B est continue en $0_E \times 0_F$.**
3. **B est bornée sur $B(0_E, 1] \times B(0_F, 1]$.**
4. **B est bornée sur le produit des sphères unité.**
5. **$\exists\, k \in \mathbb{R}^+ \quad \forall\, x, y \in E \times F \quad \|B(x,y)\| \leqslant k\,\|x\|\,\|y\|$.**

Démonstration : Les implications 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 5) sont évidentes et utilisent essentiellement des propriétés d'homogénéité. Pour montrer que 5) $\Rightarrow$ 1), on montre que, si $(x_n, y_n) \to (x, y)$ dans l'espace produit $E \times F$, on a

$$\begin{aligned}\|B(x,y) - B(x_n,y_n)\| &= \|B(x - x_n, y) - B(x_n, y_n - y)\| \\ &\leqslant k\,\|x - x_n\|\,\|y\| + k\,\|x_n\|\,\|y_n - y\|\end{aligned}$$

ce qui permet de conclure. ■

REMARQUE 7-4.19 On remarquera qu'une application bilinéaire non identiquement nulle ne peut être uniformément continue, pour la même raison qui fait que $(x, y) \to xy$ n'est pas uniformément continue sur $\mathbb{R}^2$ (pourquoi donc ?).

REMARQUE 7-4.20 La continuité du produit externe $\mathbb{K} \times E \to E$ déja vue au théorème (**7-1**.8) donne un exemple très simple de continuité d'application bilinéaire.

REMARQUE 7-4.21 On montrerait de la même manière qu'une application p-linéaire entre evn normés est continues s'il existe une majoration de la forme

$$\|B(x_1, \ldots, x_p)\| \leqslant k \prod_{i=1}^{p} \|x_i\|$$

Le théorème précédent et le théorème (**7-4**.12) donnent immédiatement

COROLLAIRE 7-4.22 **Si E, F et G sont trois espaces normés, l'application**

$$\mathcal{L}_c(E,F) \times \mathcal{L}_c(F,G) \to \mathcal{L}_c(E,G) \qquad (u,v) \mapsto v \circ u$$

est continue.

7-4.3.2 Algèbre normée

Le corollaire précédent montre en particulier que, si E est un evn, $\mathcal{L}_c(E,E) = \mathcal{L}_c(E)$ est une sous-algèbre de $(\mathcal{L}(E),+,\cdot,\circ)$. Si on la munit de la norme des applications linéaires continues, le produit $(u,v) \mapsto u \circ v$ est alors continu, avec plus précisément

$$\|u \circ v\| \leqslant \|u\|\,\|v\|$$

et également, si $E \neq \{0_E\}$, l'égalité $\|id_E\| = 1$.

Plus généralement, si un $\mathbb{K}$-ev normé A est aussi une $\mathbb{K}$-algèbre, le produit est continu ssi il existe une constante $k > 0$ telle que

$$\forall\, a,b \in A \quad \|ab\| \leqslant k\,\|a\|\,\|b\|$$

il est alors facile de voir que si l'on remplace la norme $\|\ \|$ par $k\,\|\ \|$, on aura alors la même inégalité avec $k = 1$. Ceci amène à la définition suivante :

DÉFINITION 7-4.23 (Algèbre normée) *Une $\mathbb{K}$-algèbre $(A,+,\cdot,\times,\|\ \|)$ munie d'une norme est une algèbre normée ssi*

$$\forall\, a,b \in A \quad \|ab\| \leqslant \|a\|\,\|b\|$$

*On vérifie alors aisément que l'on a forcément $\|1_A\| \geqslant 1$. On dira que l'algèbre normée est stricte (ou unitaire) ssi $\|1_A\| = 1$. On parlera d'**algèbre de Banach** si $(A,\|\ \|)$ est complet.*

Avec ce vocabulaire, les résultats précédemment obtenus peuvent s'énoncer :

THÉORÈME 7-4.24 Si E est un espace normé, l'espace $\mathcal{L}_c(E)$ muni de la norme des applications linéaires continues est une algèbre normée unitaire. Si E est un espace complet, c'est une algèbre de Banach.

THÉORÈME 7-4.25 Si X est un ensemble non vide, $(\mathcal{B}(X,\mathbb{K}),+,\cdot,\times,\|\ \|_\infty)$ est une algèbre de Banach stricte.

7-5 Compacité

Il s'agit de généraliser ici la propriété de Bolzano-Weierstrass vérifiée par les suites réelles prenant leurs valeurs dans un segment.

7-5.1 Compacts d'un espace normé

DÉFINITION 7-5.1 *Une partie K d'un espace vectoriel normé $(E,\|\ \|)$ est dite compacte si elle est non vide et si toute suite de points de K possède au moins une valeur d'adhérence **dans** K.*

On dit aussi que K possède la propriété de Bolzano-Weierstrass. Cette propriété ne change évidemment pas si on remplace la norme par une norme équivalente.

THÉORÈME 7-5.2 Tout segment est une partie compacte de $(\mathbb{R}, |\,|)$.

Par contre, $\mathbb{R}$ n'est évidemment pas compact, $[0,1[$ ne l'est également pas (toute suite de $[0,1[$ a évidemment une v.a. dans $[0,1]$, mais $u_n = 1 - \frac{1}{n}$ montre que cette v.a. n'est pas toujours dans $[0,1[$).

7-5.2 Propriétés élémentaires

THÉORÈME 7-5.3 Une réunion d'une famille finie de parties compactes de E **est compacte.**

> *Démonstration* : Exercice. On remarquera que le résultat n'est pas vrai en général pour la réunion d'une famille quelconque de compacts. ■

THÉORÈME 7-5.4 Toute partie compacte d'un espace normé est bornée.

> *Démonstration* : Si $K \subset E$ n'est pas bornée, on a
>
> $$\forall\, M \in \mathbb{R}^+ \quad \exists\, x \in K \qquad \|x\| \geqslant M$$
>
> On peut alors construire une suite de points de K telle que $\|x_n\| \geqslant n$ pour tout entier naturel n. Une telle suite ne possède évidemment pas de valeur d'adhérence. ■

THÉORÈME 7-5.5 Toute partie compacte d'un espace normé est complète, donc fermée.

> *Démonstration* : Si $(x_n)_{n\in\mathbb{N}}$ est une suite de Cauchy du compact K, elle possède une valeur d'adhérence dans K. Etant de Cauchy, elle converge vers cette valeur qui est dans K. K est bien une partie complète de E et est donc fermée. ■

THÉORÈME 7-5.6 Si A **et** B **sont deux parties non vides respectives de** E **et** F**, la partie** $A \times B$ **de l'espace normé produit** $E \times F$ **est compacte ssi** A **et** B **sont compactes.**

> *Démonstration* : Si A et B sont compacts, et si $z_n = (x_n, y_n)$ est une suite de $A\times B$, il existe une suite extraite $x'_n = x_{\varphi(n)}$ de la suite des premières coordonnées qui converge vers un point $a \in A$. De la suite de points de B de terme général $y_{\varphi(n)}$ on peut extraire une suite $y'_n = y_{\varphi\circ\psi(n)}$ qui converge vers un point $b \in B$. La suite $z_{\varphi\circ\psi(n)} = (x'_{\psi(n)}, y'_n)$ converge dans l'espace produit vers $(a,b) \in A \times B$, ce qui prouve la compacité de $A \times B$. Réciproquement, si $A \times B$ est compact, une suite $(x_n)_{n\in\mathbb{N}}$ de A peut être "remontée" en une suite (x_n, b) où b est un point fixé arbitrairement dans B. On en déduit aisément que la suite $(x_n)_{n\in\mathbb{N}}$ possède au moins une valeur d'adhérence dans A. ■

Le résultat suivant est très important. Il permet de caractériser les parties compactes "plongées" dans une partie compacte.

THÉORÈME 7-5.7 Si A est une partie compacte de $(E, \| \|)$, une partie $B \subset A$ non vide est un compact ssi B est fermée.

Démonstration : On remarquera que dans l'énoncé du théorème, le terme "fermé" signifie aussi bien fermé de $(E, \| \|)$ que fermé relatif de A (puisque A étant compact il est fermé de E, et les fermés relatifs de A sont exactement les fermés de E inclus dans A). Si B est compact, il est fermé d'après le théorème (**7-5**.5). Réciproquement, une suite de points de B est dans A compact, donc possède au moins une valeur d'adhérence dans A. Si B est fermé, cette valeur d'adhérence est forcément dans B, ce qui prouve la compacité de B. ■

7-5.3 Compacts de $\mathbb{R}$, de $\mathbb{K}^p$

Les résultats **7-5**.4 et **7-5**.5 donnent une caractérisation des parties compactes de $\mathbb{R}$ et de $(\mathbb{K}^p, \| \|_\infty)$ (et, on le verra plus loin, dans tout espace vectoriel normé de dimension finie).

THÉORÈME 7-5.8 Les parties compactes de $(\mathbb{K}^p, \| \|_\infty)$ sont exactement les parties fermées bornées (non vides).

Démonstration : Si $\emptyset \neq K$ est une partie fermée bornée de l'espace normé $(\mathbb{K}^p, \| \|_\infty)$, une suite de points de K possède au moins une valeur d'adhérence, d'après le théorème de Bolzano-Weierstrass et cette v.a. est dans K, puisque K est fermé. K est donc compacte. La réciproque a été vue dans le paragraphe précédent. ■

REMARQUE 7-5.9 Dans un espace normé de dimension infinie, les parties fermées bornées ne sont pas toutes compactes. On a vu dans le chapitre précédent (page 220) qu'on pouvait trouver dans la boule unité fermée de $(\mathcal{B}([0,1], \mathbb{R}), \| \|_\infty)$ une suite $f_n = \mathbf{1}_{]0,\frac{1}{n}]}$ qui ne possède aucune valeur d'adhérence. La boule unité fermée est fermée, bornée mais non compacte.

Attention ! Une erreur fréquemment commise est de confondre compacts de $\mathbb{R}$ et segments. La "structure" d'un compact de $\mathbb{R}$ peut être fort compliquée. Par exemple :

EXERCICE 7-5.10 Si u_n est une suite réelle convergente vers une limite l,

$$K = \{u_n \ , \ n \in \mathbb{N}\} \cup \{l\}$$

est un compact de $\mathbb{R}$. Montrer que ce résultat est valable dans un evn quelconque (attention : la caractérisation par fermé borné n'est pas correcte dans le cas général).

7-5.4 Propriété de Borel-Lebesgue

On étudie ici une caractérisation de la compacité (autre que la propriété de Bolzano-Weierstrass), qui a l'intérêt de se généraliser à des espaces plus généraux que les espaces normés et qui, même si l'énoncé semble "obscur" au départ, est d'une importance considérable pour démontrer des propriétés "globales" sur un espace compact à partir de vérifications "locales" (au voisinage de chaque point).

DÉFINITION 7-5.11 (Recouvrement ouvert) *Si $(E, \|\ \|)$ est un espace normé et A est une partie non vide de E, on appelle recouvrement ouvert de A toute famille $(\Omega_i)_{i \in I}$ d'ouverts de E, indexée par un ensemble quelconque (non vide) I, telle que*

$$A \subset \bigcup_{i \in I} \Omega_i$$

Ceci signifie que tout point de A appartient à au moins un des ouverts de la famille. On remarquera qu'alors les ensembles $\omega_i = \Omega_i \cap A$ sont des ouverts de A qui "recouvrent" A puisque leur réunion est égale à A. Par exemple, dans $(\mathbb{R}, |\ |)$, la famille $(]-\frac{1}{n}, 1-\frac{1}{n}[)_{n \in \mathbb{N}^*}$ est un recouvrement ouvert de $[0,1[$. On remarque sur cet exemple qu'une sous famille finie extraite de ce recouvrement ouvert n'est plus un recouvrement. Comme le montre le théorème qui suit, ceci est dû à la non-compacité de $[0,1[$.

THÉORÈME 7-5.12 (Borel-Lebesgue)

Borel-Lebesgue Une partie K non vide d'un espace normé est compacte ssi de tout recouvrement ouvert de K on peut extraire un sous-recouvrement fini.

Démonstration : Montrons que si K possède la propriété de Borel-Lebesgue, alors K possède la propriété de Bolzano-Weierstrass (c'est-à-dire est compact, au sens de la définition donnée plus haut). Soit $(x_n)_{n \in \mathbb{N}}$ une suite de points de K. On suppose qu'elle n'a pas de valeur d'adhérence. On a alors

$$\forall\, a \in K \quad \exists\, \varepsilon_a > 0 \quad \{n \in \mathbb{N} \mid x_n \in B(a, \varepsilon_a[\} \text{ est fini.}$$

La famille $(B(a, \varepsilon_a[)_{a \in K}$ est un recouvrement ouvert de K dont on peut, par hypothèse, extraire un sous-recouvrement fini. Soit :

$$\exists\, a_1, \dots, a_p \in K \quad K \subset \bigcup_{i=1}^{p} B(a_i, \varepsilon_{a_i}[$$

Comme chacune de ces boules ne contient qu'un nombre fini de termes de la suite x_n, on arrive évidemment à une contradiction.

Réciproquement, si K possède la propriété de Bolzano-Weierstrass, on commence par montrer que, si $\varepsilon > 0$ est fixé, il existe un nombre fini de boules ouvertes centrées en un point de K et de rayon ε qui recouvrent K. Si ce n'est pas le cas, on construit une suite $(a_n)_{n \in \mathbb{N}}$ de

points de K de la manière suivante : on choisit $a_0 \in K$ quelconque. Comme $B_0 = B(a_0, \varepsilon[$ ne recouvre pas K, on choisit $a_1 \in K - B_0$ et on pose $B_1 = B(a_1, \varepsilon[$. Par hypothèse, $B_0 \cup B_1$ ne recouvre pas K. On choisit $a_2 \in K - (B_0 \cup B_1)$ et on construit ainsi de proche en proche une suite $(a_n)_{n\in\mathbb{N}}$ de points de K vérifiant

$$\forall\, n \in \mathbb{N}^* \quad a_n \notin \bigcup_{i=1}^{n-1} B(a_i, \varepsilon[$$

On a alors, pour $n \neq m$, $\|a_n - a_m\| \geqslant \varepsilon$, ce qui est contradictoire avec l'existence d'une valeur d'adhérence pour cette suite.

On montre enfin que K possède la propriété de Borel-Lebesgue. Si $(\Omega_i)_{i\in I}$ est un recouvrement ouvert de K pour lequel il n'existe aucun sous-recouvrement fini, on montre qu'on arrive à une contradiction. Si $p \in \mathbb{N}^*$ il existe un nombre fini de boules centrées en un point de K et de rayon $\dfrac{1}{p}$ qui recouvrent K. Il existe donc forcément une de ces boules, qu'on note $B\left(x_p, \dfrac{1}{p}\right[$ telle que

$$\forall\, i \in I \quad B\left(x_p, \frac{1}{p}\right[\text{ n'est pas incluse dans } \Omega_i \qquad (1)$$

La suite (x_p) de points de K possède une valeur d'adhérence $x \in K$ et il existe un indice $i_0 \in I$ tel que $x \in \Omega_{i_0}$. Comme Ω_{i_0} est un ouvert de E, il existe $\alpha > 0$ avec $B(x, \alpha[\subset \Omega_{i_0}$. Par définition d'une valeur d'adhérence, on peut trouver des entiers aussi grands qu'on le souhaite vérifiant $\|x_p - x\| < \dfrac{\alpha}{2}$. On choisit un entier p_0 vérifiant cette propriété et tel que $\dfrac{1}{p_0} < \dfrac{\alpha}{2}$. On a alors, pour $y \in B\left(x_{p_0}, \dfrac{1}{p_0}\right[$

$$\|y - x\| \leqslant \|y - x_{p_0}\| + \|x_{p_0} - x\| < \frac{1}{p_0} + \frac{\alpha}{2} < \alpha$$

On a donc $B\left(x_{p_0}, \dfrac{1}{p_0}\right[\subset B(x, \alpha[\subset \Omega_{i_0}$, ce qui est contradictoire avec (1). K possède donc bien la propriété de Borel-Lebesgue. ■

Le cours et les exercices donneront des exemples d'utilisation de cette vision de la compacité, qu'on peut qualifier grossièrement de procédé de passage du local au global :

Soit K un compact d'un evn E et $\mathcal{P}$ une propriété pouvant être satisfaite sur certaines parties de E (on écrira alors $\mathcal{P}(A)$ pour "la propriété $\mathcal{P}$ est vérifiée sur A"). On suppose que $\mathcal{P}$ est telle que :

Si $\mathcal{P}$ est vérifiée sur A_1 et A_2, elle est vérifiée sur $A_1 \cup A_2$.
Si $A \subset B$ alors $\mathcal{P}(B) \Rightarrow \mathcal{P}(A)$.

Si tout point de K possède un voisinage (relatif à K si l'on veut) sur lequel $\mathcal{P}$ est vérifiée, alors $\mathcal{P}$ est vérifiée sur K. En effet, pour tout $x \in K$, on peut trouver un $\varepsilon_x > 0$ tel que $\mathcal{P}$ soit vérifiée sur $B(x, \varepsilon_x[\cap K$. Les $(B(x, \varepsilon_x[)_{x \in K}$ forment un recouvrement ouvert de K, dont on peut extraire un sous-recouvrement fini. Les propriétés supposées de $\mathcal{P}$ permettent ensuite de conclure. Si $f : E \to \mathbb{R}$ est une application, un exemple de propriété $\mathcal{P}$ pourrait être

$$\mathcal{P}(A) \Leftrightarrow f \text{ est bornée sur } A$$

Une fonction numérique bornée au voisinage de chaque point d'un compact K est globalement bornée sur K. On reviendra sur ce résultat dans le paragraphe suivant.

EXERCICE 7-5.13 Reprendre l'exercice (**7-5**.10)

Le passage au complémentaire permet d'obtenir une caractérisation de la compacité à l'aide des fermés :

COROLLAIRE 7-5.14 Une partie K d'un espace normé est compacte ssi pour toute famille de fermés $(F_i)_{i\in I}$ de E vérifiant

$$\left(\bigcap_{i\in I} F_i\right) \cap K = \emptyset$$

il existe une partie $J \subset I$ finie telle que

$$\left(\bigcap_{i\in J} F_i\right) \cap K = \emptyset$$

Démonstration : $\left(\bigcap_{i\in I} F_i\right) \cap K = \emptyset$ équivaut à $K \subset \bigcup_{i\in I} F_i^c$, ce qui signifie que la famille $(F_i^c)_{i\in I}$ est un recouvrement ouvert de K. Le corollaire est donc conséquence immédiate de la propriété de Borel-Lebesgue. ■

Avec la notion de fermé relatif, cela signifie qu'une famille de fermés relatifs de K qui est d'intersection vide possède une sous-famille finie qui possède la même propriété. En particulier, si $(F_n)_{n\in\mathbb{N}}$ est une suite décroissante (pour l'inclusion) de fermés non vides d'un compact (donc de compacts puisque tout fermé de F_1 est compact) $\bigcap_n F_n$ est non vide (puisque les intersections finies le sont). Si de plus le diamètre de F_n tend vers 0 lorsque n tend vers $+\infty$, on voit assez facilement que cette intersection est réduite à un point, ce qui traduit le fait que tout compact est complet.

7-5.5 Continuité et compacité

THÉORÈME 7-5.15 Si $f : E \supset K \to F$ est continue sur le compact K, alors $f(K)$ est un compact de F.

Démonstration : Si $(\Omega_i)_{i\in I}$ est un recouvrement ouvert de $f(K)$, alors pour tout $i \in I$, la partie $\omega_i = f^{-1}(\Omega_i)$ est un ouvert relatif de K, avec $\bigcup_i \omega_i = K$. On peut en extraire un sous-recouvrement fini $(\omega_i)_{i\in J}$ et on vérifie alors que $f(K) \subset \bigcup_{i\in J} \Omega_i$, ce qui prouve la compacité de $f(K)$. ■

EXERCICE 7-5.16 Démontrer ce résultat avec la propriété de Bolzano-Weierstrass.

Les conséquences de ce théorème sont multiples. On a notamment :

COROLLAIRE 7-5.17 Si $f : K \to \mathbb{R}$ est une fonction numérique continue sur le compact K, alors f est bornée sur K et atteint ses bornes, c'est-à-dire qu'il existe x et $y \in K$ tels que

$$f(x) = \inf_K f \text{ et } f(y) = \sup_K f$$

f atteint son maximum et son minimum sur K.

Démonstration : $f(K)$ est un compact, donc un fermé borné de $\mathbb{R}$. Etant borné, il admet une borne supérieure et une borne inférieure dans $\mathbb{R}$. La borne supérieure d'un sous-ensemble majoré de $\mathbb{R}$ est adhérente à ce sous-ensemble (voir la caractérisation de la borne sup.). Comme $f(K)$ est fermé, il contient tous ses points adhérents, ce qui prouve le résultat. ■

En particulier, en travaillant avec l'application continue $x \mapsto \|x - a\|$, on obtient :

COROLLAIRE 7-5.18 Si K est un compact d'un espace vectoriel normé E et $a \in E$, il existe au moins un point x_0 de K vérifiant

$$\|x_0 - a\| = d(a, K)$$

COROLLAIRE 7-5.19 Si $f : K \to F$ est continue sur le compact K, la fonction $x \mapsto \|f(x)\|$ atteint son maximum et son minimum sur K.

COROLLAIRE 7-5.20 Si K est compact, $\mathcal{C}^0(K,F)$ est fermé dans l'espace $(\mathcal{B}(K,F), \| \|_\infty)$. Il est complet si F est complet.

Démonstration : C'est le corollaire **7-3**.9 puisque

$$\mathcal{C}_b^0(K,F) = \mathcal{C}^0(K,F) \quad ■$$

Un argument de compacité donne une démonstration simple du théorème de D'Alembert :

EXERCICE 7-5.21 Si $P \in \mathbb{C}[X]$ est un polynôme non constant, montrer que

$$\lim_{|z| \to +\infty} |P(z)| = +\infty$$

En déduire l'existence d'un nombre complexe z_0 tel que

$$|P(z_0)| = \inf_{z \in \mathbb{C}} |P(z)|$$

En faisant une étude locale en z_0 montrer que $P(z_0) = 0$ (par homothétie et translation, on peut se ramener au cas où $z_0 = 0$ et $P(z_0) = 1$ si on suppose que P ne s'annule pas en z_0).

Terminons par deux théorèmes, le premier généralisant le théorème de Heine sur les fonctions continues sur un segment de $\mathbf{R}$, (résultat qui est à la base de l'intégrabilité au sens de Riemann des fonctions numériques continues sur un segment).

THÉORÈME 7-5.22 Si une fonction est continue sur un compact, elle est uniformément continue.

Démonstration : Si $f : K \to F$ était continue sans être uniformément continue, on pourrait trouver deux suites (x_n) et (y_n) de K avec $\lim_{n \to +\infty} \|x_n - y_n\| = 0$ et

$$\exists\, \varepsilon > 0 \quad \forall\, N \in \mathbb{N} \quad \exists\, n > N \quad \|f(x_n) - f(y_n)\| > \varepsilon$$

En extrayant au besoin une suite de la suite (x_n, y_n), on peut supposer que

$$\forall\, n \in \mathbb{N} \quad \|f(x_n) - f(y_n)\| > \varepsilon$$

Comme $K \times K$ est compact, on peut extraire une suite

$$(x'_n, y'_n) = (x_{\varphi(n)}, y_{\varphi(n)})$$

qui converge vers $(a, b) \in K$. Comme $\lim_{n \to +\infty} \|x'_n - y'_n\| = 0$, on a $a = b$ et par continuité de f en a,

$$\lim_{n \to +\infty} \|f(x'_n) - f(y'_n)\| = 0$$

ce qui contredit $\forall\, n \in \mathbb{N} \quad \|f(x'_n) - f(y'_n)\| > \varepsilon$. ■

THÉORÈME 7-5.23 Si $f : E \supset K \to F$ est injective et continue sur le compact K, alors la réciproque

$$f^{-1} : f(K) \to K$$

est continue. f est donc bi-(uniformément)-continue.

Démonstration : On pose $g = f^{-1}$. Pour montrer la continuité de g, il suffit de montrer que l'image réciproque d'un fermé de E est un fermé relatif de $K' = f(K)$. Si F est un fermé de E, on a $g^{-1}(F) = g^{-1}(F \cap K) = f(F \cap K)$ est l'image continue d'un compact (fermé de K), donc un compact inclus dans K' et donc un fermé de K'. ■

EXERCICE 7-5.24 Donner une démonstration de ce résultat en utilisant la propriété de Bolzano-Weierstrass.

L'exemple donné en (**7-1.7**) montre que l'hypothèse de compacité est indispensable.

7-6 Espaces normés de dimension finie

Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, nous savons que les parties compactes de $(\mathbb{K}^p, \| \|_\infty)$ sont exactement les parties fermées et bornées. Nous allons voir dans cette section des conséquences importantes de cette propriété, notamment l'équivalence des normes en dimension finie :

7-6.1 Equivalence des normes

Nous travaillerons d'abord sur l'espace $\mathbb{K}^p$, et nous nous transporterons ensuite par isomorphisme sur un espace de dimension finie sur $\mathbb{K}$.

THÉORÈME 7-6.1 Sur $\mathbb{K}^p$ toutes les normes sont équivalentes.

Démonstration : Notons $(e_i)_{1\leqslant i\leqslant p}$ la base canonique de $\mathbb{K}^p$, et montrons qu'une norme N quelconque sur $\mathbb{K}^p$ est équivalente à la norme usuelle $\| \|_\infty$. Par transitivité, on en déduira évidemment que deux normes quelconques sur $\mathbb{K}^p$ sont équivalentes.

Si $x = (x_1, \ldots, x_p) \in \mathbb{K}^p$, nous avons

$$N(x) = N\left(\sum_{i=1}^{p} x_i e_i\right) \leqslant \sum_{i=1}^{p} |x_i| \; N(e_i) \leqslant \left(\sum_{i=1}^{p} N(e_i)\right) \|x\|_\infty$$

Nous avons donc prouvé l'existence d'une constante $K > 0$ telle que

$$\forall x \in \mathbb{K}^p \quad N(x) \leqslant K \; \|x\|_\infty$$

Remarquons que cette inégalité montre que l'application

$$N : (\mathbb{K}^p, \| \|_\infty) \to \mathbb{R}$$

est K-lipschitzienne, donc continue, puisque

$$\forall x, y \in \mathbb{K}^p \quad |N(x) - N(y)| \leqslant N(x - y) \leqslant K \; \|x - y\|_\infty$$

Pour prouver le théorème, il reste à trouver une constante $k > 0$ telle que

$$\forall x \in \mathbb{K}^p \quad N(x) \geqslant k \, \|x\|_\infty$$

Par homogénéité, cela revient à prouver l'existence de $k > 0$ tel que

$$\forall x \in \mathbb{K}^p \quad \|x\|_\infty = 1 \Rightarrow N(x) \geqslant k$$

Si $\mathcal{K}$ désigne la sphère unité de $(\mathbb{K}^p, \| \, \|_\infty)$, on sait que $\mathcal{K}$ est compact (fermé borné). L'application $N : (\mathbb{K}^p, \| \, \|_\infty) \to \mathbb{R}^+$ étant continue, elle est bornée sur $\mathcal{K}$ et atteint sa borne inférieure :

$$\exists x_0 \in \mathcal{K} \quad N(x_0) = \inf_{\|x\|_\infty = 1} N(x) = k$$

Comme $x_0 \in \mathcal{K}$, on a $x_0 \neq 0_{\mathbb{K}^p}$ et donc $N(x_0) = k > 0$. On a donc bien

$$\forall x \in \mathbb{K}^p \quad k \, \|x\|_\infty \leqslant N(x) \leqslant K \, \|x\|_\infty$$

ce qui prouve l'équivalence des normes $\| \, \|_\infty$ et N. ■

COROLLAIRE 7-6.2 Sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes.

Démonstration : Si $\mathbb{E}$ est un $\mathbb{K}$-ev de dimension $p > 0$ et si $\mathcal{B} = (u_i)_{1 \leqslant i \leqslant p}$ est une base de $\mathbb{E}$, on construit un isomorphisme Φ entre $\mathbb{E}$ et $\mathbb{K}^p$ en posant

$$\Phi\left(\sum_{i=1}^{p} x_i \, u_i\right) = (x_1, \ldots, x_p)$$

Si N_1 et N_2 sont deux normes sur $\mathbb{E}$, on vérifie aisément que $N_1 \circ \Phi^{-1}$ et $N_2 \circ \Phi^{-1}$ sont des normes sur $\mathbb{K}^p$. Le théorème qui précède nous assure l'existence de deux constantes k et $K > 0$ telles que

$$\begin{aligned} \forall (x_1, \ldots, x_p) \in \mathbb{K}^p \quad & k \, N_1 \circ \Phi^{-1}(x_1, \ldots, x_p) \\ & \leqslant N_2 \circ \Phi^{-1}(x_1, \ldots, x_p) \leqslant K \, N_1 \circ \Phi^{-1}(x_1, \ldots, x_p) \end{aligned}$$

ce qui donne simplement

$$\forall x \in \mathbb{E} \quad k \, N_1(x) \leqslant N_2(x) \leqslant K \, N_1(x)$$

On a donc bien prouvé l'équivalence des normes N_1 et N_2. ■

Sur un $\mathbb{K}$-ev E de dimension p, il n'y a donc qu'une seule topologie d'espace normé : si $\mathcal{B} = (u_i)_{1 \leqslant i \leqslant p}$ est une base de E, toutes les normes sur E sont équivalentes à la norme $\| \, \|_{\infty, \mathcal{B}}$ définie par

$$\left\| \sum_{i=1}^{p} x_i \, u_i \right\|_{\infty, \mathcal{B}} = \max_{1 \leqslant i \leqslant p} |x_i|$$

La convergence d'une suite de vecteurs de E pour une norme arbitraire est donc la convergence coordonnée par coordonnée dans une base quelconque de E.

COROLLAIRE 7-6.3 Tout espace vectoriel normé de dimension finie est complet.

Démonstration : Si $(E, \| \|)$ est un espace normé de dimension p et $(y_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de cet espace, cette suite est aussi de Cauchy pour la norme $\| \|_{\infty,\mathcal{B}}$ définie plus haut. Si on décompose dans la base $\mathcal{B}$

$$y_n = \sum_{i=1}^{p} y_n^i u_i$$

chacune des suites $(y_n^i)_{n\in\mathbb{N}}$ est de Cauchy dans $(\mathbb{K}, | |)$, donc converge vers une limite $y^i \in \mathbb{K}$. On a donc

$$\lim_{n\to+\infty} y_n = \sum_{i=1}^{p} y^i u_i$$

pour la norme $\| \|_{\infty,\mathcal{B}}$, donc aussi pour $\| \|$ qui lui est équivalente. ■

COROLLAIRE 7-6.4 Les compacts d'un espace normé de dimension finie sont exactement les parties fermées bornées.

Démonstration : Soit $(E, \| \|)$ un espace de dimension finie. Nous savons qu'un compact de E est nécessairement fermé borné. Réciproquement, si $\mathcal{K} \neq \emptyset$ est une partie fermée bornée de E, c'est aussi une partie fermée bornée pour la norme équivalente $\| \|_{\infty,\mathcal{B}}$ définie précédemment. L'application

$$\psi : (\mathbb{K}^p, \| \|_\infty) \to \left(\mathbb{E}, \| \|_{\infty,\mathcal{B}}\right) \quad (x_1, \ldots, x_p) \mapsto \sum_{i=1}^{p} x_i u_i$$

est une bijection qui conserve les distances, donc un homéomorphisme. L'ensemble $\psi^{-1}(\mathcal{K})$ est ainsi une partie fermée bornée de $(\mathbb{K}^p, \| \|_\infty)$, donc un compact de cet espace. $\mathcal{K} = \psi\left(\psi^{-1}(\mathcal{K})\right)$ est donc compact de E, comme image continue d'un compact. ■

EXERCICE 7-6.5 Soit $(E, \| \|)$ un espace normé et u une forme linéaire sur E. Montrer que l'application

$$N : E \to \mathbb{R}^+ \quad x \mapsto N(x) = |u(x)| + \|x\|$$

est une norme sur E. Montrer que N est équivalente à $\| \|$ si et seulement si u est continue. En déduire qu'un espace vectoriel normé $(E, \| \|)$ est de dimension finie si et seulement si toutes les normes sur E sont équivalentes à $\| \|$.

Indication : il suffit, si E est de dimension infinie, de montrer qu'on peut trouver une forme linéaire non continue sur E. On pourra admettre que tout sous-espace vectoriel de E possède un supplémentaire.

Si $(E, \| \|)$ est un espace normé et $F \subset E$ est un sous-espace **de dimension finie**, la restriction de la norme à F munit ce sous-espace d'une structure d'espace normé. Il est remarquable que, si on munit E d'une autre norme, non nécessairement équivalente à $\| \|$, la topologie induite par cette autre norme sur F sera la même que celle induite par $\| \|$, puisque les restrictions des normes à F seront toutes équivalentes.

COROLLAIRE 7-6.6 Si F est un sous-espace de dimension finie d'un espace normé $(E, \| \|)$, F est un fermé de E.

> *Démonstration* : $(F, \| \|)$ est un espace normé de dimension finie, et est donc complet. La partie F de E est complète donc fermée. ■

7-6.2 Continuité des applications linéaires

THÉORÈME 7-6.7 Si $(E, \| \|)$ est un espace normé de dimension finie, et $(F, \| \|)$ est un autre espace normé (non nécessairement de dimension finie), toute application linéaire de E dans F est continue.

> *Démonstration* : Soit u une application linéaire de E vers F. Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant p}$ est une base de E, on a pour tout $x \in E$ décomposé dans la base $\mathcal{B}$
>
> $$\|u(x)\| = \left\| u\left(\sum_{i=1}^{p} x_i\, e_i\right)\right\| \leqslant \sum_{i=1}^{p} |x_i|\ \|u(e_i)\| \leqslant \left(\sum_{i=1}^{p} \|u(e_i)\|\right) \|x\|_{\infty,\mathcal{B}}$$
>
> Cette inégalité prouve la continuité de u, puisque $\| \|_{\infty,\mathcal{B}}$ est équivalente à toute norme sur E. ■

Si $(E, \| \|)$ est un espace normé de dimension finie et $(F, \| \|)$ un e.v.n. quelconque, on a donc

$$\mathcal{L}_c(E, F) = \mathcal{L}(E, F)$$

On notera bien que l'hypothèse de dimension finie ne porte que sur l'espace de départ.

EXERCICE 7-6.8 On suppose que E et F sont comme précédemment, et $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant p}$ est une base de E. On munit l'espace $\mathcal{L}(E, F)$ de la norme des applications linéaires continues

$$\forall u \in \mathcal{L}(E, F) \quad \|u\| = \sup_{\substack{x \in E \\ x \neq 0_E}} \|u(x)\|$$

Montrer que

$$N(u) = \sum_{i=1}^{p} \|u(e_i)\|$$

définit une autre norme sur $\mathcal{L}(E, F)$, et que cette norme est équivalente à la norme des applications linéaires continues. En déduire que, lorsque l'espace E de départ est de dimension finie, la convergence en norme dans $\mathcal{L}(E, F)$ n'est pas différente de la convergence simple.

REMARQUE 7-6.9 Si E_1 et E_2 sont deux espaces normés de dimension finie, rapportés à des bases respectives $\mathcal{B}_1 = (e_j)_{1\leqslant j\leqslant p}$ et $\mathcal{B}_2 = (f_k)_{1\leqslant k\leqslant q}$, on montre que toute application bilinéaire de $E_1 \times E_2$ dans $(F, \| \|)$ est continue : si φ est une telle application, on a en effet, avec des notations évidentes

$$\|\varphi(x,y)\| = \left\|\varphi\left(\sum_{j=1}^{p} x_j\, e_j, \sum_{k=1}^{q} y_k\, f_k\right)\right\| \leqslant \left[\sum_{\substack{1\leqslant j\leqslant p\\ 1\leqslant k\leqslant q}} \|\varphi(e_j, f_k)\|\right] \|x\|_{\infty,\mathcal{B}_1}\, \|y\|_{\infty,\mathcal{B}_2}$$

Cette inégalité prouve bien la continuité de φ. Ce résultat se généralise évidemment aux applications multilinéaires définies sur un produit d'espaces normés de dimensions finies.

7-6.3 Norme matricielle subordonnée à des normes sur $\mathbb{K}^n$ et $\mathbb{K}^p$

Si p et n sont des entiers naturels non nuls, l'espace $\mathcal{M}_{p,n}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel de dimension np sur lequel toutes les normes sont équivalentes. Si A appartient à $\mathcal{M}_{p,n}(\mathbb{K})$, le morphisme canoniquement associé à A est l'application linéaire

$$\Phi_A : \mathbb{K}^n \to \mathbb{K}^p \quad X \mapsto AX$$

Comme $\mathbb{K}^n$ est de dimension finie, Φ_A est continue pour tout choix de normes sur $\mathbb{K}^n$ et $\mathbb{K}^p$. Si N_1 et N_2 sont des normes respectivement sur $\mathbb{K}^n$ et $\mathbb{K}^p$, la norme de Φ_A considérée comme application linéaire continue entre $(\mathbb{K}^n, N_1)$ et $(\mathbb{K}^p, N_2)$ est

$$\|\Phi_A\| = \sup_{\substack{X\in\mathbb{K}^n\\ X\neq 0}} \frac{N_2(AX)}{N_1(X)}$$

Comme la correspondance $A \mapsto \Phi_A$ est bijective (et linéaire), ceci nous amène à la définition :

DÉFINITION 7-6.10 *Si N_1 et N_2 sont des normes respectivement sur $\mathbb{K}^n$ et $\mathbb{K}^p$, la norme sur $\mathcal{M}_{p,n}(\mathbb{K})$ subordonnée à N_1 et N_2 est définie par*

$$\forall A \in \mathcal{M}_{p,n}(\mathbb{K}) \quad \|A\| = \sup_{\substack{X\in\mathbb{K}^n\\ X\neq 0}} \frac{N_2(AX)}{N_1(X)}$$

Il s'agit donc d'une norme[1] sur $\mathcal{M}_{p,n}(\mathbb{K})$ vérifiant

$$\forall X \in \mathbb{K}^n \quad N_2(AX) \leqslant \|A\|\, N_1(X)$$

[1]Comme la sphère unité de $(\mathbb{K}^n, N_1)$ est compacte et que l'application $X \mapsto N_2(AX)$ est continue sur cette sphère, on pourra toujours trouver un vecteur $X_0 \in \mathbb{K}^n$ avec $N_1(X_0) = 1$ et tel que

$$N_2(AX_0) = \max_{\substack{X\in\mathbb{K}^n\\ N_1(X)=1}} N_2(AX) = \sup_{\substack{X\in\mathbb{K}^n\\ X\neq 0}} \frac{N_2(AX)}{N_1(X)}$$

et nous savons que $\|A\|$ est la plus petite constante vérifiant cette inégalité pour tout X de $\mathbb{K}^n$. Comme nous travaillons avec des normes d'applications linéaires continues, nous avons aussi :

PROPOSITION 7-6.11 Si N_1, N_2 et N_3 sont des normes respectives sur $\mathbb{K}^n$, $\mathbb{K}^p$ et $\mathbb{K}^m$, nous aurons

$$\forall A \in \mathcal{M}_{p,n}(\mathbb{K}) \quad \forall B \in \mathcal{M}_{m,p}(\mathbb{K}) \quad \|BA\| \leqslant \|B\| \, \|A\|$$

où la norme de A est subordonnée à N_1 et N_2, celle de B est subordonnée à N_2 et N_3 et celle de BA est subordonnée à N_1 et N_3.

En particulier, si $\| \ \|$ est une norme sur $\mathbb{K}^n$, elle permettra de définir sur $\mathcal{M}_n(\mathbb{K})$ une norme qui lui est subordonnée, que nous noterons encore par $\| \ \|$

$$\forall A \in \mathcal{M}_n(\mathbb{K}) \quad \|A\| = \sup_{\substack{X \in \mathbb{K}^n \\ X \neq 0}} \frac{\|AX\|}{\|X\|}$$

On aura alors

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}) \quad \|AB\| \leqslant \|A\| \, \|B\| \quad \text{et} \quad \|I_n\| = 1$$

et donc $(\mathcal{M}_n(\mathbb{K}), \| \ \|)$ est une algèbre normée (complète). Utiliser une telle norme sur $\mathcal{M}_n(\mathbb{K})$ est souvent commode, notamment parce qu'elle vérifie

$$\forall A \in \mathcal{M}_n(\mathbb{K}) \quad \forall p \in \mathbb{N}^* \quad \|A^p\| \leqslant \|A\|^p$$

EXEMPLE 7-6.12 Déterminons la norme sur $\mathcal{M}_n(\mathbb{C})$ subordonnée à la norme $\| \ \|_\infty$ sur $\mathbb{C}^n$:

Si $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$, et X est le vecteur colonne de composantes $(x_i)_{1 \leqslant i \leqslant n}$, nous avons

$$AX = Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

avec

$$\forall i \quad y_i = \sum_{j=1}^{n} a_{ij} \, x_j$$

ce qui donne

$$\forall i \quad |y_i| \leqslant \sum_{j=1}^{n} |a_{ij}| \, |x_j| \leqslant \left(\sum_{j=1}^{n} |a_{ij}| \right) \|X\|_\infty$$

On en déduit

$$(*) \qquad \|AX\|_\infty = \max_{1 \leqslant i \leqslant n} |y_i| \leqslant \left[\max_{1 \leqslant i \leqslant n} \left(\sum_{j=1}^{n} |a_{ij}| \right) \right] \|X\|_\infty$$

Par définition de $\|A\|$, nous aurons

$$\|A\| \leqslant \max_{1\leqslant i\leqslant n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

Mais il est facile de trouver un vecteur $X \neq 0_{\mathbb{K}^n}$ pour lequel l'inégalité $(*)$ est en fait une égalité : il suffit de choisir un indice i_0 pour lequel

$$\sum_{j=1}^{n}|a_{i_0 j}| = \max_{1\leqslant i\leqslant n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

et de choisir pour vecteur X un vecteur de la forme

$$X = \begin{pmatrix} e^{i\alpha_1} \\ \vdots \\ e^{i\alpha_n} \end{pmatrix}$$

avec $\alpha_k \in \mathbb{R}$ (et donc $\|X\|_\infty = 1$) et tel que

$$\forall j \quad a_{i_0 j}\, x_j = a_{i_0 j}\, e^{i\alpha_j} \in \mathbb{R}^+$$

puisqu'alors l'inégalité triangulaire

$$\left|\sum_{j=1}^{n} a_{i_0 j}\, x_j\right| \leqslant \left(\sum_{j=1}^{n}|a_{i_0 j}|\right)\|X\|_\infty$$

est une égalité. On a donc finalement

$$\|A\| = \max_{1\leqslant i\leqslant n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

EXERCICE 7-6.13 Montrer que la norme sur $\mathcal{M}_n(\mathbb{C})$ subordonnée à la norme $\|\ \|_1$ est donnée, pour $A = (a_{ij})$ par la formule

$$\|A\| = \max_{1\leqslant j\leqslant n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$$

On voit qu'il s'agit d'une formule analogue à celle obtenue précédemment, mais on a simplement remplacé la matrice A par sa transposée. Ceci pourrait se voir aussi en remarquant que l'espace $(\mathbb{C}^n, \|\ \|_1)$ est le dual topologique de $(\mathbb{C}^n, \|\ \|_\infty)$, ceci signifiant simplement que la norme $\|\ \|_1$ d'un vecteur ligne est la norme de la forme linéaire continue définie par cette ligne sur $(\mathbb{K}^n, \|\ \|_\infty)$.

Si $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$, il est souvent peu commode de déterminer explicitement $\|A\|$ pour une norme subordonnée à une norme sur $\mathbb{K}^n$. Par exemple, dans le cas réel, nous verrons, au chapitre sur les espaces euclidiens, que la norme subordonnée à $\|\ \|_2$ est donnée par

$$\|A\| = \sqrt{\text{la plus grande valeur propre de } {}^tAA}$$

alors qu'il n'est nullement évident, *a priori*, que cette formule définisse une norme sur $\mathcal{M}_n(\mathbb{R})$.

EXERCICE 7-6.14 Si $\lambda \in \mathbb{K}$ est une valeur propre de $A \in \mathcal{M}_n(\mathbb{K})$, montrer que, pour toute norme subordonnée à une norme sur $\mathbb{K}^n$, on a

$$|\lambda| \leqslant \|A\|$$

EXERCICE 7-6.15 Si $A \in \mathcal{M}_n(\mathbb{C})$, on appelle **rayon spectral de** A et on note $\rho(A)$ le réel positif

$$\rho(A) = \max\{|\lambda|, \quad \lambda \text{ valeur propre de } A\}$$

L'exercice précédent nous montre que, pour toute norme subordonnée à une norme sur $\mathbb{C}^n$, nous avons

$$\rho(A) \leqslant \|A\|$$

Le but de cet exercice est de démontrer la formule, dite "du rayon spectral"

$$(*) \quad \rho(A) = \lim_{k \to +\infty} \left\|A^k\right\|^{\frac{1}{k}}$$

quelle que soit la norme utilisée sur $\mathcal{M}_n(\mathbb{C})$.

- En utilisant l'équivalence des normes sur $\mathcal{M}_n(\mathbb{C})$, montrer que, si la propriété $(*)$ est vérifiée pour une norme particulière sur $\mathcal{M}_n(\mathbb{C})$, elle l'est pour toute norme.

- Si on suppose $\rho(A) = 0$, que peut-on en déduire pour la matrice A ? Montrer alors que $(*)$ est une évidence.

- Si on choisit une norme subordonnée à une norme sur $\mathbb{C}^n$, montrer que la plus petite valeur d'adhérence m (dans $\mathbb{R}^+ \cup \{+\infty\}$) de la suite $\left(\|A^k\|^{\frac{1}{k}}\right)_{k \in \mathbb{N}^*}$ vérifie

$$\rho(A) \leqslant m$$

- Soit $A \in \mathcal{M}_n(\mathbb{C})$, on peut trouver une matrice $P \in \mathcal{GL}_n(\mathbb{C})$ telle que $P^{-1}AP$ soit triangulaire supérieure. Si $p \in \mathbb{N}^*$, on considère la matrice diagonale

$$D_p = \operatorname{Diag}\left(1, \frac{1}{2^p}, \frac{1}{3^p}, \dots, \frac{1}{n^p}\right)$$

  Que peut-on dire des coefficients diagonaux de la matrice

$$A_p = (PD_p)^{-1} A (PD_p)$$

  et des coefficients situés au-dessus de la diagonale principale, si p est "grand" ?

- Montrer que, si $\varepsilon > 0$ est donné, on peut trouver une norme N sur $\mathcal{M}_n(\mathbb{C})$ subordonnée à une norme sur $\mathbb{C}^n$ telle que

$$\rho(A) \leqslant N(A) \leqslant \rho(A) + \varepsilon$$

- Conclure.

REMARQUE 7-6.16 Une norme N sur $\mathcal{M}_n(\mathbb{K})$ est dite **norme matricielle**[2] si et seulement si elle vérifie

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}) \quad N(AB) \leqslant N(A)\, N(B)$$

Les normes subordonnées aux normes sur $\mathbb{K}^n$ vérifient cette propriété. Mais on peut trouver des normes matricielles qui ne proviennent pas de normes sur $\mathbb{K}^n$. Vérifier par exemple que

$$\|A\| = \sqrt{\operatorname{trace}({}^t\!AA)}$$

est une norme matricielle sur $\mathcal{M}_n(\mathbb{R})$ (expliciter simplement $\|A\|$ en fonction des coefficients (a_{ij}) de A), qui ne provient pas d'une norme sur $\mathbb{R}^n$. Cette norme est très utilisée en pratique, car elle se calcule facilement et vérifie notamment

$$\forall p \in \mathbb{N}^* \quad \|A^p\| \leqslant \|A\|^p$$

On a de plus (exercice), pour cette norme

$$\forall X \in \mathbb{R}^n \quad \|AX\|_2 \leqslant \|A\|\ \|X\|_2$$

En fait, toute norme sur $\mathcal{M}_n(\mathbb{K})$ peut être transformée, par multiplication par un scalaire, en un norme matricielle : si N est une norme quelconque, la continuité du produit se traduit par l'existence d'une constante $k > 0$ telle que

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}) \quad N(AB) \leqslant k\, N(A)\, N(B)$$

ce qui signifie que $k\,N$ est une norme matricielle.

7-6.4 Complément : théorème de Riesz

Nous avons vu que, dans un espace normé de dimension finie, les parties compactes sont les parties fermées et bornées. En particulier la boule unité fermée est compacte. Nous nous proposons ici de voir que cette propriété caractérise les espaces de dimension finie.

7-6.4.1 Projection sur un sous-espace de dimension finie

Dans un espace normé $(E, \|\ \|)$, considérons un sous-espace vectoriel F de dimension finie. Si $x \in E$, on essaie d'approcher au mieux x par un élément de F. C'est une démarche très souvent utilisé en analyse : les éléments de F sont souvent des objets "simples", possédant des propriétés intéressantes[3], qu'on utilise pour approcher un élément quelconque de E, plus "compliqué".

On considère donc

$$d = d(x, F) = \inf_{y \in F} \|x - y\|$$

[2] On précise parfois "sous-multiplicative".

[3] Par exemple, si E est l'espace de toutes les fonctions continues sur $[0,1]$ à valeurs dans $\mathbb{C}$, muni de la norme de la convergence uniforme, F peut être l'espace des fonctions polynômes de degré inférieur à un entier p fixé.

la distance de x à F, et on cherche un élément $y_0 \in F$ tel que $\|x - y_0\| = d(x, F)$. Remarquons que l'existence d'un tel y_0 n'est nullement évidente *a priori*, puisque la borne inférieure d'un ensemble de réels minoré n'est pas en général élément de cet ensemble.

Nous savons déjà que F est fermé dans $(E, \| \|)$, et donc le problème est évidemment résolu si $d = 0$, puisque

$$d(x, F) = 0 \Leftrightarrow x \in \overline{F} = F$$

et dans ce cas la solution (unique) du problème est évidemment $y_0 = x$. Nous supposerons donc $d > 0$, avec $x \notin F$.

Nous cherchons en fait à minimiser la fonction continue

$$\varphi : F \to \mathbb{R}^+ \quad y \mapsto \|x - y\|$$

Un argument souvent utilisé pour montrer qu'une fonction continue atteint sa borne inférieure est la compacité. Mais ici, le domaine de définition de φ est F qui n'est pas compact (sauf dans le cas $F = \{0_E\}$) puisque non borné. On peut cependant se ramener à travailler sur une partie compacte de F, puisqu'il est inutile de chercher y_0 "trop loin" :

Considérons en effet la boule fermée de centre x et de rayon $d + 1$, et sa trace sur F

$$\mathcal{K} = F \cap B(x, d + 1]$$

Par définition de d, nous savons que $\mathcal{K} \neq \emptyset$. $\mathcal{K}$ est une partie fermée de F (intersection de F avec un fermé de E), évidemment bornée. C'est donc un compact de l'espace F de **dimension finie**. L'application

$$\varphi|_{\mathcal{K}} : \mathcal{K} \to \mathbb{R}^+$$

atteint sa borne inférieure sur $\mathcal{K}$ et donc

$$\exists\, y_0 \in \mathcal{K} \quad \|x - y_0\| = \inf_{y \in \mathcal{K}} \|x - y\|$$

Par construction, nous avons

$$\|x - y_0\| \leqslant d + 1$$

et par conséquent, en distinguant les points de F qui sont dans $\mathcal{K}$ et ceux qui n'y sont pas,

$$\forall y \in F \quad \|x - y_0\| \leqslant \|x - y\|$$

Nous avons donc obtenu :

PROPOSITION 7-6.17 Si F est un sous-espace de dimension finie d'un espace normé $(E, \| \|)$, tout élément $x \in E$ possède une meilleure approximation par un élément de F, c'est à dire

$$\exists\, y_0 \in F \quad \|x - y_0\| = d(x, F)$$

Un tel y_0 est une projection[4] de x sur F.

[4]Attention à la terminologie : cette notion n'a en général aucun rapport avec celle de projecteur sur un sous-espace parallèlement à un supplémentaire. Nous verrons cependant que les deux notions sont liées dans le cadre très particulier des espaces préhilbertiens,

REMARQUE 7-6.18 Un élément $x \in E$ peut posséder plusieurs projections sur F. Ce phénomène est lié à la géométrie des boules de l'espace E : par exemple, sur l'espace $(\mathbb{R}^2, \| \|_\infty)$, si F est le sous-espace $\mathbb{R} \times \{0\}$, quelles sont les projections de $x = (0, 1)$ sur F ? (Indication : faire un dessin). Même question avec $F = \mathbb{R}.u$, avec $u = (1, 1)$.

Lorsque l'espace normé est un espace préhilbertien (c'est à dire lorsque la norme provient d'un produit scalaire), nous verrons (section **14-2**) qu'il y a unicité de la projection.

7-6.4.2 Théorème de Riesz

THÉORÈME 7-6.19 Un espace normé $(E, \| \|)$ est de dimension finie si et seulement si sa boule unité fermée est compacte.

Par homothétie et translation, on voit aussi que, si une boule fermée de E (de rayon $r > 0$) est compacte, alors E est de dimension finie.

Indication pour la démonstration : on montre en fait que, si E est de dimension infinie, on peut construire une suite de vecteurs $(e_n)_{n \in \mathbb{N}^*}$, tous de norme égale à 1 et vérifiant

$$\forall n \neq m \quad \|e_n - e_m\| \geqslant 1$$

Montrer que, si une telle suite existe la boule unité fermée de E n'est pas compacte. Pour construire une telle suite, on opère de proche en proche : on choisit e_1 avec $\|e_1\| = 1$; si on suppose ensuite construits $(e_i)_{1 \leqslant i \leqslant n}$, on considère le sous espace

$$F_n = \text{vect}\,(e_i)_{1 \leqslant i \leqslant n} \subsetneq E$$

Si $x \in E - \text{vect}\,(e_i)_{1 \leqslant i \leqslant n}$, il possède une projection y sur F_n. Considérer alors le vecteur

$$z = \frac{x - y}{\|x - y\|}$$

7-7 Connexité par arcs

Il s'agit d'une notion qui va nous permettre de généraliser le théorème des valeurs intermédiaires pour les fonctions réelles définies sur un intervalle de $\mathbb{R}$:

7-7.1 Théorème des valeurs intermédiaires : fonctions définies sur un intervalle

THÉORÈME 7-7.1 Si I est un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ est continue, alors $f(I)$ est un intervalle.

Démonstration : Les intervalles sont les parties convexes de $\mathbb{R}$. Il suffit donc de montrer que, si α et β sont dans $f(I)$ avec $\alpha < \beta$, alors $[\alpha, \beta] \subset f(I)$. On peut trouver a et $b \in I$ avec $\alpha = f(a)$ et $\beta = f(b)$.

Sans nuire à la généralité, nous supposerons $a < b$. Pour montrer que $[\alpha, \beta] \subset f(I)$, nous prenons $\gamma \in]\alpha, \beta[$ arbitraire et nous montrons qu'il existe un $c \in]a, b[$ avec $f(c) = \gamma$. Pour cela considérons

$$X = \{x \in [a, b] \mid f(x) \leqslant \gamma\}$$

Cet ensemble est non vide (il contient a), et est majoré par b. Il possède donc une borne supérieure c qui est limite d'une suite de points de X. Par continuité, on a

$$f(c) \leqslant \gamma$$

De plus, comme $f(b) = \beta > \gamma$, il existe un $\alpha > 0$ tel que la fonction f ne prenne que des valeurs strictement supérieures à γ sur $[b - \alpha, b]$ (continuité de f en b). On a donc $c < b$ et

$$\forall x \in]c, b] \quad f(x) > \gamma$$

Par continuité de f en c, on aura

$$f(c) \geqslant \gamma$$

soit finalement $f(c) = \gamma$. ■

Le théorème qui précède entraîne notamment :

COROLLAIRE 7-7.2 Si I est un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ est continue, s'il existe a et $b \in I$ avec

$$f(a) f(b) < 0$$

alors f s'annule au moins une fois entre a et b.

COROLLAIRE 7-7.3 Si $f : I \to \mathbb{R}$ est continue et ne s'annule pas, alors $f(I) \subset \mathbb{R}^{*+}$ ou $f(I) \subset \mathbb{R}^{*-}$. En d'autres termes, f a un signe constant sur l'intervalle I.

Comme les segments sont les intervalles compacts de $\mathbb{R}$, nous avons aussi :

COROLLAIRE 7-7.4 Si K est un segment de $\mathbb{R}$ et $f : K \to \mathbb{R}$ est continue, alors $f(K)$ est un segment.

7-7.2 Connexité par arcs

7-7.2.1 Chemin continu

DÉFINITION 7-7.5 *Si $(E, \| \|)$ est un espace normé, on appelle chemin continu dans E toute application continue*

$$\gamma : [0, 1] \to E$$

L'ensemble image

$$\gamma([0, 1]) = \{\gamma(t),\ t \in [0, 1]\}$$

est appelé support du chemin γ. C'est une partie compacte de E.

DÉFINITION 7-7.6 *Si γ est un chemin continu dans E, les points*

$$x_0 = \gamma(0) \ \text{et}\ x_1 = \gamma(1)$$

sont appelés respectivement origine et extrémité du chemin γ.

On utilise le segment $[0,1]$ comme ensemble de variation du paramètre t pour des raisons de commodité : si a et b sont réels avec $a < b$, toute application continue $f : [a,b] \to E$ permet de définir un chemin continu γ par

$$\forall t \in [0,1] \quad \gamma(t) = f(a + t(b-a))$$

DÉFINITION 7-7.7 *Si γ_1 et γ_2 sont deux chemins continus dans E tels que l'extrémité de γ_1 soit égal à l'origine de γ_2*

$$\gamma_1(1) = \gamma_2(0)$$

l'application $\gamma : [0,1] \to E$ définie par

$$\begin{cases} \gamma(t) = \gamma_1(2t) \ \text{pour}\ t \in \left[0, \frac{1}{2}\right] \\ \gamma(t) = \gamma_2(2t-1) \ \text{pour}\ t \in \left[\frac{1}{2}, 0\right] \end{cases}$$

est un chemin continu, qu'on appelle chemin juxtaposé de γ_1 et γ_2. On le note

$$\gamma = \gamma_1 \dot{+} \gamma_2$$

7-7.2.2 Partie connexe par arcs

Une partie d'un espace normé connexe par arcs sera, d'une certaine manière, "en un seul morceau" :

DÉFINITION 7-7.8 *Une partie A d'un espace normé est dite connexe par arcs si et seulement si, pour tous points a et $b \in A$, il existe un chemin continu tracé dans A (c'est-à-dire dont le support est inclus dans A) d'origine a et d'extrémité b.*

EXEMPLE 7-7.9 En particulier, toute partie convexe d'un espace normé est évidemment connexe par arcs. Dans un espace normé de dimension différente de 1, le complémentaire d'un point est une partie connexe par arcs (exercice) non convexe.

Le théorème des valeurs intermédiaires **7-7**.1 permet de caractériser les parties connexes par arcs de $\mathbb{R}$:

PROPOSITION 7-7.10 Les parties de $\mathbb{R}$ connexes par arcs sont les intervalles.

Démonstration : Tout intervalle est évidemment connexe par arcs. Réciproquement, si A est un connexe par arcs de $\mathbb{R}$, et si a et $b \in A$ avec $a < b$, il existe un chemin continu tracé dans A

$$\gamma : [0,1] \to A$$

avec $\gamma(0) = a$ et $\gamma(1) = b$. Le théorème des valeurs intermédiaires appliqué à γ entraîne

$$[a,b] \subset \gamma([0,1]) \subset A$$

ce qui montre que A est une partie convexe de $\mathbb{R}$, donc un intervalle. ■

EXERCICE 7-7.11 Soit $(A_i)_{i\in I}$ une famille de parties connexes par arcs d'un espace normé telles que

$$\forall i,j \in I \quad A_i \cap A_j \neq \emptyset$$

Montrer que la réunion des $(A_i)_{i\in I}$ est connexe par arcs.

7-7.2.3 Connexité par arcs et continuité

Le résultat qui suit est la généralisation du théorème **7-7**.1 :

PROPOSITION 7-7.12 Si $(E, \| \|)$ et $(F, \| \|)$ sont deux espaces normés, A une partie connexe par arcs de E et

$$f : A \to F$$

une fonction continue, alors $f(A)$ est une partie connexe par arcs de F.

Démonstration : Soient $x, y \in f(A)$. Il existe a et $b \in A$ avec

$$x = f(a) \text{ et } y = f(b)$$

Comme A est connexe par arcs, on peut trouver un chemin continu γ tracé dans A avec

$$\gamma(0) = a \text{ et } \gamma(1) = b$$

L'application continue $f \circ \gamma$ est alors un chemin continu tracé dans $f(A)$ reliant x à y. ■

COROLLAIRE 7-7.13 Si A est une partie connexe par arcs d'un espace normé et $f : A \to \mathbb{R}$ est une application continue, alors $f(A)$ est un intervalle.

L'exercice qui suit est une belle application de ce résultat :

EXERCICE 7-7.14 Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une application dérivable. Montrer que $f'(I)$ est un intervalle. Ce résultat, appelé théorème de Darboux, montre qu'une dérivée, même lorsqu'elle n'est pas continue, vérifie toujours le théorème des valeurs intermédiaires.

Indication : On considère

$$D = \{(x, y) \in I \times I \mid x < y\}$$

Montrer que D est une partie connexe par arcs de $\mathbb{R}^2$. On définit

$$\varphi : D \to \mathbb{R} \quad (x, y) \mapsto \frac{f(x) - f(y)}{x - y}$$

Montrer que $\varphi(D)$ est un intervalle J et comparer J avec $f'(I)$ (on pourra utiliser le théorème **8-5**.4).

REMARQUE 7-7.15 La notion de connexité par arcs peut ainsi être utilisée pour prouver que deux parties d'espaces normés ne sont pas homéomorphes : nous savons par exemple que

$$U = \{z \in \mathbb{C} \mid |z| = 1\}$$

et $[0, 1[$ ne sont pas homéomorphes, puisque U est compact et $[0, 1[$ ne l'est pas. De même, il ne peut exister d'homéomorphime

$$\varphi : [0, 1] \to U$$

L'argument de compacité n'est plus utilisable. Mais, si φ existait, il induirait un homéomorphisme de $[0, 1] - \left\{\frac{1}{2}\right\}$ vers $U - \left\{\varphi\left(\frac{1}{2}\right)\right\}$. Ce n'est pas possible, puisque ce dernier ensemble est connexe par arcs (faire un dessin) alors que $[0, 1] - \left\{\frac{1}{2}\right\}$ ne l'est pas. On peut de même voir (exercice) qu'il ne peut y avoir d'homéomorphisme entre $]0, 1[$ et $[0, 1[$. Voir aussi à ce sujet la remarque suivant le théorème **8-1**.9.

7-7.2.4 Partition d'un connexe par arcs en ouverts relatifs

PROPOSITION 7-7.16 Si A est une partie connexe par arcs d'un espace normé $(E, \| \|)$, il n'existe pas de partition de A en deux ouverts relatifs non vides.

Démonstration : Supposons qu'une telle partition existe. On peut alors écrire

$$A = A_1 \cup A_2$$

avec $A_1 \neq \emptyset$, $A_2 \neq \emptyset$ tels qu'il existe deux ouverts Ω_1 et Ω_2 de E avec

$$A_1 = \Omega_1 \cap A,\ A_2 = \Omega_2 \cap A \text{ et } \Omega_1 \cap \Omega_2 \cap A = \emptyset$$

Choisissons $a_1 \in A_1$, $a_2 \in A_2$ et un chemin continu γ tracé dans A et reliant a_1 à a_2. Si nous notons 1_{A_1} la fonction caractéristique de A_1 (qui vaut 1 sur A_1 et 0 sur son complémentaire), l'application

$$f : [0, 1] \to \mathbb{R} \quad f = 1_{A_1} \circ \gamma$$

est continue : elle ne prend que les valeurs 0 et 1 et, si $t_0 \in [0,1]$ vérifie $f(t_0) = 0$ (par exemple), cela signifie que $\gamma(t_0) \in A_2 = \Omega_2 \cap A$; comme A_2 est un ouvert relatif de A et comme γ est continue, il existe alors un voisinage V de t_0 dans $[0,1]$ tel que

$$\forall t \in V \quad \gamma(t) \in A_2$$

ce qui donne $f(t) = f(t_0)$ pour tout t de V et prouve la continuité de f en t_0. La fonction f est continue avec $f(0) = 1$ et $f(1) = 0$, ce qui contredit le théorème des valeurs intermédiaires, puisque

$$f([0,1]) = \{0,1\}$$

On ne peut donc partitionner A en deux ouverts relatifs non vides. ■

Comme le complémentaire dans A d'un ouvert relatif de A est un fermé relatif de A, une partition de A en deux ouverts relatifs est aussi une partition en deux fermés relatifs. Nous avons dons aussi :

COROLLAIRE 7-7.17 Si A est une partie connexe par arcs d'un espace normé $(E, \| \|)$, il n'existe pas de partition de A en deux fermés relatifs non vides. Les seuls sous-ensembles de A à la fois ouverts et fermés relatifs de A sont donc $\emptyset$ et A.

REMARQUE 7-7.18 Une partie A d'un espace normé est dite connexe si elle possède ces propriétés. La connexité par arcs est donc un cas particulier de la propriété plus générale de connexité.

Un espace normé est évidemment connexe par arcs ($t \mapsto (1-t)a + tb$ relie a à b). On a donc :

COROLLAIRE 7-7.19 Les seuls ouverts-fermés d'un espace normé $(E, \| \|)$ sont $\emptyset$ et E.

Le résultat qui suit est souvent dénommé "théorème du passage des douanes" :

PROPOSITION 7-7.20 Soit A une partie d'un espace normé, a un point intérieur à A et b un point intérieur au complémentaire de A. Si γ est un chemin continu reliant a à b, il existe $t_0 \in [0,1]$ avec

$$\gamma(t_0) \in \mathrm{Fr}(A)$$

Démonstration : Si ce n'était pas le cas, en notant B le complémentaire de A, le connexe par arcs $\gamma([0,1])$ s'écrirait

$$\gamma([0,1]) = \left(\mathring{A} \cap \gamma([0,1])\right) \cup \left(\mathring{B} \cap \gamma([0,1])\right)$$

ce qui en donnerait une partition en deux ouverts non vides. ■

La proposition **7-7**.16 est souvent utilisée de la manière suivante ("raisonnements par connexité") : pour montrer que tous les points d'un connexe par arcs A vérifient une propriété $\mathcal{P}$, on montre que

$$\{a \in A \mid a \text{ vérifie la propriété } \mathcal{P}\}$$

est non vide et est à la fois un ouvert et un fermé relatif de A :

Par exemple, une application $f : A \to X$ où A est une partie d'un espace normé et X est un ensemble quelconque est dite **localement constante** si et seulement si

$$\forall a \in A \quad \exists V \text{ voisinage de } a \quad f \text{ est constante sur } V \cap A$$

Lorsque A est connexe par arcs, une telle fonction est constante sur A :

PROPOSITION 7-7.21 Si A est connexe par arcs, toute application localement constante définie sur A est constante.

Démonstration : Si f est une telle application et $a \in A$ est fixé, considérons l'ensemble X des $x \in A$ possédant un voisinage V_x tel que

$$\forall y \in V_x \cap A \quad f(y) = f(a)$$

Comme f est constante au voisinage de a, X contient a, et est donc non vide. Par sa définition, il est clair que X est un ouvert relatif de A. Pour conclure, il suffit de montrer que X est un fermé[5] de A : si $(x_n)_{n\in\mathbb{N}}$ est une suite de points de X convergente vers $x \in A$, montrons que $x \in X$. Par hypothèse, il existe un voisinage V_x de x tel que f soit constante égale à $f(x)$ sur $V_x \cap A$. Comme la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x, il existe un entier n_0 tel que

$$n \geqslant n_0 \Rightarrow x_n \in V_x \cap A$$

On a donc $f(a) = f(x_n) = f(x)$, et donc $x \in X$. ■

EXERCICE 7-7.22 Soit $\mathcal{O}$ un ouvert de $\mathbb{R}^2$ connexe par arcs. Montrer que deux points quelconques de $\mathcal{O}$ peuvent être reliés par une ligne polygonale continue incluse dans $\mathcal{O}$, constituée de segments parallèles à l'un ou l'autre des axes de coordonnées.

Indication : fixer $a \in \mathcal{O}$, et considérer

$$X_a = \{b \in \mathcal{O} \mid b \text{ peut être relié à } a \text{ par un tel chemin}\}$$

[5]Si X était une partie d'un espace normé, il suffirait de dire que f est continue puisque localement constante, et que

$$X = f^{-1}(\{f(a)\})$$

est un fermé de A comme image réciproque d'un fermé.

7-7.2.5 Complément : composantes connexes par arcs

EXERCICE 7-7.23 Soit X une partie non vide d'un espace normé. On définit dans X une relation binaire par

$$a \mathcal{R} b \Leftrightarrow \text{il existe un chemin continu dans } X \text{ reliant } a \text{ et } b$$

Montrer que $\mathcal{R}$ est une relation d'équivalence dans X. Les différentes classes d'équivalence réalisent une partition de X. Montrer que ces classes d'équivalence sont des parties connexes par arcs de X, et que la classe d'équivalence de $a \in X$ est la plus grande partie de X connexe par arcs (pour l'inclusion) et contenant a. Une telle partie de X est appelée composante connexe par arcs de X.

EXERCICE 7-7.24 Trouver les composantes connexes par arcs de l'ouvert de $\mathbb{R}^2$

$$\mathcal{O} = \{(x, y) \in \mathbb{R}^2 \mid xy \neq 1\}$$

EXERCICE 7-7.25 Soient x et $y \in \mathbb{R}$. Posons

$$z = \arctan x + \arctan y$$

Montrer que

$$z = \pm\frac{\pi}{2} \Leftrightarrow xy = 1$$

Si $\mathcal{O}$ est l'ouvert de $\mathbb{R}^2$ défini dans l'exercice précédent, montrer que

$$\varphi : \mathcal{O} \to \mathbb{R} \quad (x, y) \mapsto \arctan x + \arctan y - \arctan \frac{x + y}{1 - xy}$$

est continue, et prend ses valeurs dans $\pi\mathbb{Z}$. En déduire que φ est constante sur chaque composante connexe par arcs de $\mathcal{O}$. Trouver la valeur de cette constante pour chacune de ces composantes.

EXERCICE 7-7.26 Montrer que les composantes connexes par arcs d'un ouvert non vide de $\mathbb{R}$ sont des intervalles **ouverts**. En déduire que tout ouvert de $\mathbb{R}$ est réunion d'une famille d'intervalles ouverts disjoints deux à deux. Lorsque cette famille n'est pas finie, on peut montrer qu'elle peut être indexée par une partie infinie de $\mathbb{Q}$ (choisir dans chacun des intervalles un nombre rationnel). Comme une partie infinie de $\mathbb{Q}$ peut être mise en bijection avec $\mathbb{N}$ (voir la section **9-5.1**), on a ainsi montré que tout ouvert de $\mathbb{R}$ est réunion d'une famille finie ou d'une suite d'intervalles ouverts deux à deux disjoints.

7-8 Exercices

EXERCICE 7-8.1 Soit $(E, \| \ \|)$ un espace normé, A et B deux parties non vides de E telles que $\bar{A} \cap B = \bar{B} \cap A = \emptyset$. Montrer qu'il existe deux ouverts U et V tels que $A \subset U$, $B \subset V$ et $U \cap V = \emptyset$. Montrer qu'en général cette propriété est fausse si on suppose seulement $A \cap B = \emptyset$. On pourra considérer les applications $x \mapsto d(x, A)$ et $x \mapsto d(x, B)$.

EXERCICE 7-8.2 Soient A et B deux parties non vides d'un evn. On note

$$A + B = \{a + b \mid a \in A, b \in B\}$$

Montrer que :

1. A ou B ouvert $\Rightarrow \quad A+B$ ouvert.

2. A fermé et B compact $\Rightarrow \quad A+B$ fermé. Que se passe-t-il si B est seulement fermé?

3. A et B compacts $\Rightarrow \quad A+B$ compact.

EXERCICE 7-8.3 Dans $\mathcal{M}_n(\mathbb{C})$ muni d'une norme quelconque, on considère

$$D_2 = \{A \in \mathcal{M}_n(\mathbb{C}) \mid A \text{ diagonalisable}\}$$

$$D_1 = \{A \in \mathcal{M}_n(\mathbb{C}) \mid A \text{ possède } n \text{ valeurs propres distinctes}\}$$

1. Montrer que D_1 est un ouvert dense dans $\mathcal{M}_n(\mathbb{C})$. On pourra utiliser le fait que, pour $P, Q \in \mathbb{C}[X]$ non nuls :

$$P \wedge Q \neq 1 \iff$$
$$\exists U, V \in \mathbb{C}[X] \quad \deg(U) < \deg(Q), \quad \deg(V) < \deg(P), \quad UP + VQ = 0$$

Montrer que cette propriété peut se traduire sur les coefficients de P et Q par la nullité d'un déterminant (le "résultant" de P et Q)

2. Montrer que D_2 n'est ni ouvert ni fermé. Préciser son intérieur et son adhérence.

EXERCICE 7-8.4 Soient A et B deux parties fermées d'un espace normé. Montrer qu'en général il n'existe pas $(a, b) \in A \times B$ avec $d(a, b) = d(A, B)$. Que se passe-t-il si une des parties est compacte? Plus généralement, si A est une partie d'un espace normé $(E, \|\ \|)$, on note $P(A)$ la propriété :

$$P(A) \ : \quad \forall B \text{ fermé } \neq \emptyset \quad (\ A \cap B = \emptyset \ \Rightarrow \ d(A, B) > 0\ )$$

Montrer que A compact $\Rightarrow P(A)$ et que, si $P(A)$ est vérifiée, alors A est fermé et la frontière de A est compacte.

EXERCICE 7-8.5 <u>Théorème du point fixe avec paramètre</u> :
Soient X et Λ deux parties d'espaces normés avec X complet et $f : X \times \Lambda \to X$ telle que :

$$\exists k \in [0, 1[\ \forall \lambda \in \Lambda \ f(\bullet, \lambda) \text{ soit } k\text{-contractante}$$

$$\forall x \in X \ f(x, \bullet) \text{ est continue } : \Lambda \to X$$

Montrer que, $\forall \lambda \in \Lambda$ l'équation $f(x, \lambda) = x$ possède une unique solution x_λ et que $\lambda \mapsto x_\lambda$ est continue.

EXERCICE 7-8.6 Soit X une partie compacte d'un espace normé et $f : X \to X$ telle que :

$$\forall x, y \ x \neq y \Rightarrow d(f(x), f(y)) < d(x, y)$$

1. Montrer que f n'est pas nécessairement une contraction.

2. Montrer que f possède un unique point fixe (considérer $\varphi(x) = d(f(x), x)$).

3. Montrer que, $\forall a \in X$ la suite $(f^{(n)}(a))_{n \in \mathbb{N}}$ converge vers ce point fixe.

4. Si X est un convexe compact d'un evn et f vérifie $d(f(x), f(y)) \leqslant d(x,y)$ montrer que f possède au moins un point fixe, et que l'hypothèse de convexité est indispensable.

EXERCICE 7-8.7 Soit $\varphi \in E = \mathcal{C}^0([0,1], \mathbb{R})$. Montrer que l'application

$$\Psi \ : \ (E, \| \ \|_\infty) \to (\mathbb{R}, | \ |)$$

définie par : $\Psi(f) = \displaystyle\int_0^1 \varphi(t) f(t) dt$ est continue et calculer sa norme.

EXERCICE 7-8.8 Soit K un espace compact, F un K-ev normé. On munit $\mathcal{C}^0(K, F)$ de la norme de la convergence uniforme $\|f\|_\infty = \sup_{x \in K} \|f(x)\|$. Démontrer l'équivalence des propositions, (g_n) étant une suite de $\mathcal{C}^0(K, F)$:

1. g_n converge uniformément vers g sur K.

2. $\forall x \in K \ \ \forall (x_n)_n \in K^{\mathbb{N}} \ \ x_n \to x \ \ \Rightarrow \ \ g(x_n) \to g(x)$

EXERCICE 7-8.9 Soit X un compact d'un espace normé et L_c l'ensemble des applications c-lipchitziennes de X dans $\mathbb{R}$. Montrer que si (f_n) est une suite d'éléments de L_c convergeant simplement vers f, alors la convergence est uniforme.

EXERCICE 7-8.10 Soit $E = \mathcal{C}^0([0,1], \mathbb{R})$ muni de $\| \ \|_\infty$. On pose

$$N_k = \{f \in E \mid f \text{ est } k - \text{lipschitzienne}\}$$

pour $k \geqslant 0$ et $N = \bigcup_{k \geqslant 0} N_k$. Montrer que N_k est fermé dans E, que $\overset{\circ}{N} = \emptyset$ et $\overline{N} = E$. Sur N on définit $\|f\| = \sup_{x \neq y} \dfrac{|f(x) - f(y)|}{|x - y|}$. Montrer que ce n'est pas une norme, mais que sa restriction à $N_0 = \{f \in N \mid f(0) = 0\}$ en est une. Est-elle équivalente à $\| \ \|_\infty$?

EXERCICE 7-8.11 Pour $K = \mathbb{R}$ ou $\mathbb{C}$, $E = \mathcal{M}_n(K)$ est muni d'une norme quelconque.

1. $K = \mathbb{R}$ et $F = \{M \in E \mid M^2 = I_n\}$. F est-il fermé ? compact ? Quels sont les points d'accumulation de F (c'est à dire les limites de suites injectives et convergentes de points de F) ?

2. Déterminer l'adhérence de $\{A \in \mathcal{M}_2(\mathbb{C}) \mid \exists p \in \mathbb{N} \ \ A^p = I_2\}$.

EXERCICE 7-8.12 Soit E un $\mathbb{R}$-ev normé de dimension finie et $f \in L(E)$ tel que la suite $(f^p)_{p \in \mathbb{N}}$ soit bornée. On note $g_p = \dfrac{1}{p} \displaystyle\sum_{i=0}^{p-1} f^i$. Montrer que la suite $(g_p)_{p \in \mathbb{N}}$ converge vers un projecteur.

EXERCICE 7-8.13 Montrer que l'application $f \ : \ \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$f(x,y) = \Big(4x - \sin(x+y), 3y - 2\mathrm{arctg}(x-y)\Big)$$

est bijective. Sa réciproque est elle continue ?

EXERCICE 7-8.14 On munit $E = \mathcal{C}^0([0,1],\mathbb{R})$ de la norme de la convergence uniforme. Soit $\varphi : E \to \mathbb{R}$ définie par $\varphi(f) = \int_0^{\frac{1}{2}} f(t)dt - \int_{\frac{1}{2}}^1 f(t)dt$. Montrer que φ est continue et déterminer sa norme. Cette norme est-elle atteinte sur la sphère unité de E ?

EXERCICE 7-8.15 Soit $E = \mathcal{C}^0([0,1],\mathbb{R})$ et F un sev de E de dimension finie. Montrer qu'une suite de fonctions de F convergeant simplement sur $[0,1]$ converge en fait uniformément, et que la limite est dans F.

EXERCICE 7-8.16 Soit $E = \{x \in \mathbb{R}^n \mid \lim\limits_{+\infty} x_n = 0\}$, muni de $\| \|_\infty$. Soit $\lambda \in \mathbb{R}^\mathbb{N}$ telle que $\forall n \in \mathbb{N}\ \lambda_n \neq 0$ et $U_\lambda : E \to \mathbb{R}^\mathbb{N}$ définie par $x \mapsto (\lambda_n x_n)$. Donner une CNS sur λ pour que $U_\lambda \in \mathcal{L}(E)$. Etudier alors la continuité de U_λ et déterminer $\|U_\lambda\|$. Trouver une CNS pour que U_λ soit un homéomorphisme de E.

Chapitre 8

Fonctions d'une variable réelle

Pour l'essentiel, nous étudierons dans ce chapitre des fonctions définies sur un intervalle $I \subset \mathbb{R}$ à valeurs dans un $\mathbb{K}$-espace normé $(E, \| \|)$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Si $E = \mathbb{R}$ ou $\mathbb{C}$, on parlera de fonctions numériques, c'est un cas évidemment très important. Enfin, dans certains paragraphes (notamment la section **8-3**), on pourra considérer des fonctions définies sur une partie d'un espace normé.

8-1 Fonctions monotones sur un intervalle

Qui dit fonction monotone dit fonction **réelle**! On considère donc ici des fonctions $I \to \mathbb{R}$ où I est un intervalle de $\mathbb{R}$ non réduit à un point.

8-1.1 Définitions

Une fonction $f : I \to \mathbb{R}$ est dite croissante sur I si et seulement si

$$\forall x, y \in I \qquad x \leqslant y \Rightarrow f(x) \leqslant f(y)$$

Il s'agit donc ici de fonction croissante au sens large. Si $x < y \Rightarrow f(x) < f(y)$ la fonction f est dite strictement croissante. Si $-f$ est (strictement) croissante, la fonction f est (strictement) décroissante[1]. Une fonction est dite monotone sur I si elle est croissante ou décroissante sur cet intervalle. Enfin, une fonction est dite monotone par morceaux sur un segment $[a, b]$ s'il existe une subdivision

$$a = x_0 < x_1 < \cdots < x_n = b$$

[1]Dans la terminologie anglo-saxonne, on dit plutôt "croissante" pour "strictement croissante" et "non décroissante" pour "croissante au sens large".

de $[a, b]$ telle que

$$f|_{]x_i, x_{i+1}[} \text{ soit monotone pour } i = 0, \dots, n-1$$

Les fonctions "usuelles" (faut-il dire les "bonnes fonctions" ?) sont monotones par morceaux sur tout segment. Il faudrait bien se garder de croire qu'il s'agisse là du comportement "générique" des fonctions définies sur un segment. Prenons un exemple :

> Si $x_0 \in I$, on rappelle qu'une fonction $f : I \to \mathbb{R}$ présente un minimum local en x_0 s'il existe $\alpha > 0$ tel que
>
> $$\forall x \in]x_0 - \alpha, x_0 + \alpha[\cap I \qquad f(x) \geqslant f(x_0)$$
>
> (le maximum local est strict si l'inégalité est stricte pour $x \neq x_0$). Cette propriété est évidemment vérifiée s'il existe $\alpha > 0$ tel que f soit décroissante sur $]x_0 - \alpha, x_0] \cap I$ et croissante sur $[x_0, x_0 + \alpha[\cap I$. Cette situation n'est cependant pas nécessaire pour avoir un minimum local en x_0 comme le montre l'exemple
>
> $$f(x) = x^2 \cos^2 \frac{1}{x} + x^4 \sin^2 \frac{1}{x} \text{ prolongée par continuité par } f(0) = 0$$
>
> qui présente un minimum absolu en 0, mais n'est monotone sur aucun intervalle d'extrémité 0 (pourquoi ?).

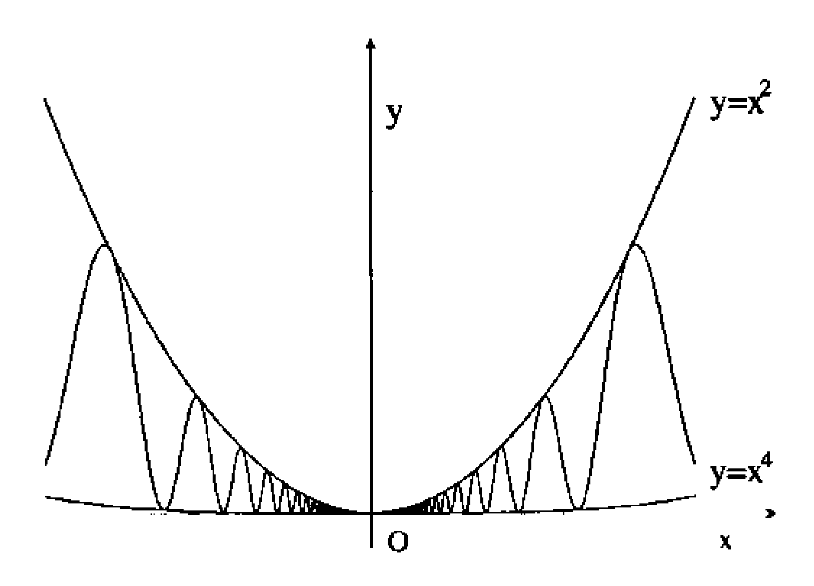


Figure **8.1** – *Fonction non monotone par morceaux admettant un minimum absolu*

Par des manipulations élémentaires d'inégalités, on démontre facilement le théorème :

THÉORÈME 8-1.1 Si f et $g : I \to \mathbb{R}$ sont monotones et $\lambda \in \mathbb{R}$

$$\lambda > 0 \Rightarrow \lambda f \textbf{ monotone de même sens que } f.$$
$$f \textbf{ et } g \textbf{ de même sens} \Rightarrow f + g \textbf{ monotone (de même sens que } f \textbf{ et } g\textbf{)}$$
$$f \textbf{ et } g \textbf{ de même sens et } > 0 \Rightarrow fg \textbf{ monotone (de même sens que } f \textbf{ et } g\textbf{)}$$
$$f \textbf{ monotone de signe constant sur } I \Rightarrow \frac{1}{f} \textbf{ monotone (de sens opposé)}$$

Si $f : I \to \mathbb{R}$ et $g : J \to \mathbb{R}$ sont monotones avec $f(I) \subset J$

$$f \textbf{ et } g \textbf{ de même sens} \Rightarrow g \circ f \textbf{ croissante}$$
$$f \textbf{ et } g \textbf{ de sens opposés} \Rightarrow g \circ f \textbf{ décroissante}$$

Ces résultats permettent parfois d'obtenir des résultats de monotonie en évitant le recours systématique à une étude de signe de dérivée (si les fonctions étudiées sont dérivables ...)

EXEMPLE 8-1.2 Etudier les variations de la fonction définie par

$$f(x) = \frac{1}{\cos^4 x + 2\cos^3 x + 3\cos^2 x + 2\cos x + 1}$$

On remarque que cette fonction est paire, 2π-périodique. On étudie donc ses variations sur $[0, \pi]$. La fonction $x \mapsto \cos x$ est décroissante $[0, \pi] \to [-1, 1]$. Il suffit donc étudier les variations sur $[-1, 1]$ de

$$u \mapsto \frac{1}{u^4 + 2u^3 + 3u^2 + 2u + 1} = \frac{1}{(u^2 + u + 1)^2}$$

La fonction polynôme $u \mapsto u^2+u+1$ est > 0, décroissante sur $[-1, -\frac{1}{2}]$, croissante sur $\left[-\frac{1}{2}, 1\right]$. Son inverse au carré possède un sens de variation opposé, et comme pour $x \in [0, \pi]$, $\cos x = -\dfrac{1}{2}$ ssi $x = \dfrac{2\pi}{3}$, les résultats précédemment évoqués donnent aisément le tableau de variation :

x	0		$\frac{2\pi}{3}$		π
$f(x)$	$\frac{1}{9}$	$\nearrow$	$\frac{16}{9}$	$\searrow$	1

8-1.2 Propriétés de continuité et monotonie

Nous énoncerons les résultats dans le cas des fonctions croissantes. Ils se modifient de manière évidente dans le cas d'une fonction décroissante.

THÉORÈME 8-1.3 Soient I un intervalle de $\mathbb{R}$ ouvert en son extrémité droite, $f : I \to \mathbb{R}$ croissante et $\beta = \sup I \in \mathbb{R} \cup \{+\infty\}$. On a alors

$$\lim_{\substack{x \to \beta \\ <}} f(x) = \sup\{f(x),\ x \in I\} = \sup_I f \in \mathbb{R} \cup \{+\infty\}$$

En particulier, f est majorée sur I si et seulement si f possède une limite finie à gauche en β. De même si I est ouvert à son extrémité gauche α, on a

$$\lim_{\substack{x \to \alpha \\ >}} f(x) = \inf \{f(x),\ x \in I\} = \inf_I f \in \mathbb{R} \cup \{-\infty\}$$

Démonstration : Supposons $I = (\alpha, \beta[$ et supposons f majorée sur I. Soit $M = \sup_I f \in \mathbb{R}$. Par définition de la borne supérieure, pour $\varepsilon > 0$ on peut trouver $c \in I$ avec $M - \varepsilon < f(c) \leqslant M$. Comme f est croissante on a aussi

$$\forall x \in [c, \beta[\qquad M - \varepsilon < f(x) \leqslant M$$

ce qui prouve que $\lim_{\substack{x \to \beta \\ <}} f(x) = M$. Si f n'est pas majorée, pour $A \in \mathbb{R}$ quelconque, on trouvera de même $c \in I$ avec

$$\forall x \in [c, \beta[\qquad A \leqslant f(x)$$

ce qui donnera bien $\lim_{\substack{x \to \beta \\ <}} f(x) = +\infty$. L'étude en l'extrémité droite de I serait analogue. ■

Remarquons que si $I = (\alpha, \beta]$ contient sa borne supérieure et si f est croissante sur I, $f(\beta)$ est un majorant de f, donc de $f|_{(\alpha,\beta[}$, et en conséquence f possède une limite finie à gauche en β vérifiant $f(\beta^-) \leqslant f(\beta)$. Plus généralement

COROLLAIRE 8-1.4 Si f est monotone sur un intervalle I, f possède en tout point intérieur à I une limite à droite et une limite à gauche (une limite à droite en l'extrémité gauche de l'intervalle et/ou une limite à gauche en une extrémité droite lorsque ces extrémités sont dans I). Si $a < b < c$ sont trois points de I et si f est croissante, on a

$$f(a^+) \leqslant f(b^-) \leqslant f(b) \leqslant f(b^+) \leqslant f(c^-)$$

Démonstration : Supposons par exemple f croissante sur I. Si b est un point intérieur à I, il suffit d'appliquer le théorème précédent à $f|_{I \cap]-\infty, b[}$ (majorée par $f(b)$) et à $f|_{I \cap]b, +\infty[}$ (minorée par $f(b)$) pour obtenir l'existence de limites à droite et à gauche vérifiant

$$f(b^-) \leqslant f(b) \leqslant f(b^+)$$

Si ensuite $a < b$, $f(b^-)$ est un majorant de $f|_{]a,b[}$, et vérifie donc

$$f(b^-) \geqslant \inf f|_{]a,b[} = f(a^+)$$

et on montre de même que, pour $b < c$, $f(b^+) \leqslant f(c^-)$. ■

EXERCICE 8-1.5 Montrer que l'ensemble des points de I où f présente une discontinuité est fini ou dénombrable (c'est-à-dire peut être mis en bijection avec $\mathbb{N}$). Indication : on sait (voir par exemple le chapitre sur les familles sommables et les séries) que la réunion d'une suite d'ensembles finis est finie ou dénombrable. On commencera par montrer que, pour $[a,b] \subset I$ et $p \in \mathbb{N}^*$, l'ensemble

$$\left\{x \in [a,b] \mid f\left(x^+\right) - f\left(x^-\right) \geqslant \frac{1}{p}\right\}$$

est fini.

COROLLAIRE 8-1.6 Si $f : I \to \mathbb{R}$ est monotone, f est continue sur I si et seulement si $f(I)$ est intervalle.

Démonstration : Si f est continue, le théorème des valeurs intermédiaires montre que $f(I)$ est un intervalle. Réciproquement, supposons par exemple f croissante discontinue en un point x_0 intérieur à I (le cas d'une extrémité se traiterait de la même façon). On a alors, par exemple,

$$f(x_0) < f\left(x_0^+\right)$$

Si $y > x_0$ est dans I, on a $\{f(x_0), f(y)\} \subset f(I)$, et cependant $[f(x_0), f(y)] \not\subset f(I)$, puisque tout $\alpha \in]f(x_0), f\left(x_0^+\right)[$ est dans l'intervalle $[f(x_0), f(y)]$ mais ne peut appartenir à $f(I)$. ■

REMARQUE 8-1.7 Si f est continue strictement croissante sur un intervalle I, le théorème des valeurs intermédiaires montre alors facilement que

$$\begin{array}{lll}
I = [\alpha,\beta] & (\alpha,\beta \in \mathbb{R}) & \Rightarrow\ f(I) = [f(\alpha), f(\beta)] \\
I = [\alpha,\beta[& (\alpha \in \mathbb{R}, \beta \in \mathbb{R} \cup \{+\infty\}) & \Rightarrow\ f(I) = [f(\alpha), \lim\limits_{\beta^-} f[\\
I =]\alpha,\beta] & (\beta \in \mathbb{R}, \alpha \in \mathbb{R} \cup \{-\infty\}) & \Rightarrow\ f(I) =]\lim\limits_{\alpha^+} f, f(\beta)] \\
I =]\alpha,\beta[& (\alpha \in \mathbb{R} \cup \{-\infty\}, \beta \in \mathbb{R} \cup \{+\infty\}) & \Rightarrow\ f(I) =]\lim\limits_{\alpha^+} f, \lim\limits_{\beta^-} f[
\end{array}$$

8-1.3 Homéomorphismes d'intervalles

Sur un intervalle, l'injectivité d'une fonction réelle continue se traduit par une monotonie stricte :

THÉORÈME 8-1.8 Soit $f : I \to \mathbb{R}$ continue. f est injective si et seulement si elle est strictement monotone.

Démonstration : Si f est strictement monotone, elle est clairement injective. Réciproquement, si f est injective, considérons

$$\mathcal{D} = \{(x,y) \in I \times I \mid x < y\}$$

Il est clair que $\mathcal{D}$ est une partie convexe (donc connexe par arcs) de $\mathbb{R}^2$. L'application

$$\varphi : \mathcal{D} \to \mathbb{R} \qquad (x,y) \mapsto \varphi(x,y) = f(y) - f(x)$$

est continue, donc $\varphi(\mathcal{D})$ est un intervalle de $\mathbb{R}$. L'injectivité de f montre que cet intervalle ne contient pas 0, et est donc inclus dans $\mathbb{R}^{*+}$ ou $\mathbb{R}^{*-}$. f est alors strictement croissante ou décroissante. ■

On déduit du théorème précédent qu'un homéomorphisme $f : I \to J$ entre deux intervalles est forcément strictement monotone. Réciproquement, une fonction continue strictement monotone sur un intervalle permet de construire un homéomorphisme d'intervalles :

THÉORÈME 8-1.9 Si $f : I \to \mathbb{R}$ est continue strictement monotone, l'image $J = f(I)$ est intervalle et f induit un homéomorphisme entre I et J.

Démonstration : J est un intervalle d'après le théorème des valeurs intermédiaires. Comme f est strictement monotone elle est injective et réalise donc une bijection $I \to J$. Sa bijection réciproque f^{-1} est évidemment strictement monotone de même sens que f. La continuité de f^{-1} résulte du corollaire **8-1**.6 puisque $f^{-1}(J) = I$ est un intervalle. ■

Compte tenu de la remarque **8-1**.7 et du fait qu'il est très facile à l'aide d'homothéties, de translations et d'inversions de construire des homéomorphismes entre intervalles de même type bornés ou non (par exemple $x \mapsto \dfrac{1}{x-1} + 3$ réalise un homéomorphisme entre $[0, 1[$ et $]-\infty, 2]$, $x \mapsto \dfrac{1}{x} + \dfrac{1}{x-1}$ entre $]0, 1[$ et $\mathbb{R}$), la relation "être homéomorphe à" réalise une partition de l'ensemble des intervalles de $\mathbb{R}$ en cinq classes : $\{\emptyset\}$, l'ensemble des singletons, l'ensemble des segments non réduits à un point, l'ensemble des intervalles semi-ouverts (contenant une seule extrémité) et l'ensemble des intervalles ouverts aux deux extrémités.

8-1.4 Rappel : fonctions circulaires réciproques

8-1.4.1 Fonction arcsin

La fonction continue $x \mapsto \sin x$ **restreinte à** $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ est stritement croissante, et vérifie $\sin\left(\pm\frac{\pi}{2}\right) = \pm 1$. Elle réalise donc un homéomorphisme de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ dans $[-1, 1]$. On appelle fonction arcsin l'homéomorphisme réciproque :

$$\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

avec, pour $x \in [-1, 1]$,

$$\alpha = \arcsin x \Leftrightarrow \left(x = \sin\alpha \quad \text{et} \quad \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$$

La fonction arcsin est clairement impaire, strictement croissante.

Attention aux simplifications! S'il est clair que

$$\forall x \in [-1, 1] \quad \sin(\arcsin x) = x,$$

pour $\alpha \in \mathbb{R}$, par contre $\beta = \arcsin(\sin\alpha)$ est l'unique réel de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ qui vérifie $\sin\alpha = \sin\beta$. Il n'est donc égal à α que si $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Dans le cas général

$$\exists k \in \mathbb{Z} \quad \beta = \alpha + 2k\pi \text{ ou } \beta = \pi - \alpha + 2k\pi$$

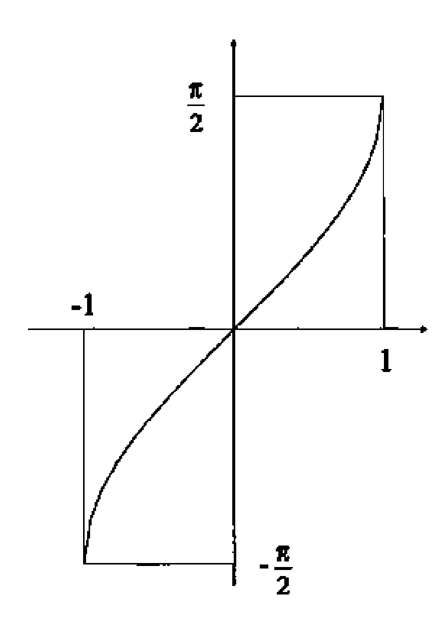


Figure **8.2** – *Fonction arcsinus*

EXERCICE 8-1.10 Simplifier, pour $x \in]0,1[$, l'expression

$$\varphi(x) = \arcsin\left(2x\sqrt{1-x^2}\right)$$

8-1.4.2 Fonction arccos

La fonction continue $x \mapsto \cos x$ **restreinte à** $[0,\pi]$ est stritement décroissante, et vérifie $\cos(0) = 1$ et $\cos(\pi) = -1$. Elle réalise donc un homéomorphisme de $[0,\pi]$ dans $[-1,1]$. On appelle fonction arccos l'homéomorphisme réciproque :

$$\arccos : [-1,1] \to [0,\pi]$$

avec, pour $x \in [-1,1]$

$$\alpha = \arccos x \Leftrightarrow (x = \cos\alpha \quad \text{et} \quad \alpha \in [0,\pi])$$

La fonction arccos est strictement **décroissante**. Elle vérifie

$$\forall x \in [-1,1] \quad \arccos(-x) = \pi - \arccos x$$

Ici encore, s'il est clair que

$$\forall x \in [-1,1] \quad \cos(\arccos x) = x,$$

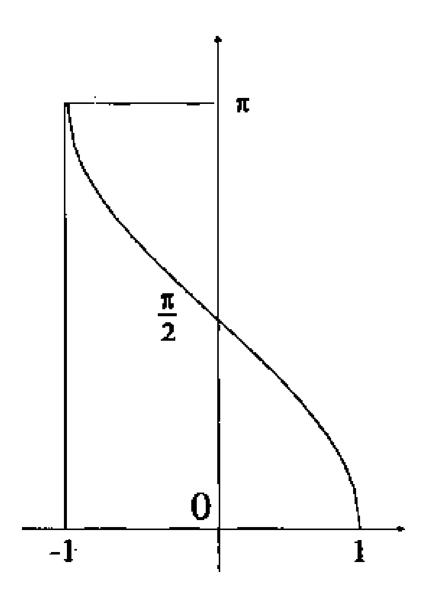


Figure **8.3** – *Fonction arccosinus*

pour $\alpha \in \mathbf{R}$, le réel $\beta = \arccos(\cos\alpha)$ est l'unique réel de $[0, \pi]$ qui vérifie $\cos\alpha = \cos\beta$. Il n'est donc égal à α que si $\alpha \in [0, \pi]$. Dans le cas général

$$\exists\, k \in \mathbb{Z} \quad \beta = \pm\alpha + 2k\pi$$

La formule $\sin\left(\frac{\pi}{2} - t\right) = \cos t$ pour $t \in \mathbf{R}$ donne immédiatement (exercice)

$$\forall\, x \in [-1, 1] \quad \arcsin x + \arccos x = \frac{\pi}{2}$$

8-1.4.3 Fonction arctan

La fonction $\tan : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \mathbf{R}$ est continue, strictement croissante et tend vers $\pm\infty$ aux bornes de l'intervalle. Elle réalise donc un homéomorphisme de l'intervalle $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ dans $\mathbf{R}$, et on appelle arctan l'homéomorphisme réciproque. La fonction arctan est strictement **croissante, impaire.** Elle vérifie clairement

$$\forall\, x \in \mathbf{R}^* \quad \arctan x + \arctan\frac{1}{x} = \pm\frac{\pi}{2}$$

le signe $\pm$ correspondant à celui de x. Cette formule est importante, car elle permet notamment d'effectuer une étude asymptotique de la fonction arctan au voisinage de $\pm\infty$ à partir de l'étude de la même fonction au voisinage de 0. C'est une conséquence immédiate (exercice) de la formule reliant $\tan\left(\alpha \pm \frac{\pi}{2}\right)$ à $\frac{1}{\tan\alpha}$. Ici encore, s'il est clair que

$$\forall\, x \in \mathbf{R} \quad \tan(\arctan x) = x,$$

pour $\alpha \in \mathbf{R}$, le nombre $\beta = \arctan(\tan\alpha)$ est l'unique réel de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ qui vérifie $\tan\alpha = \tan\beta$. Il n'est donc égal à α que si $\alpha \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Dans le cas général

$$\exists\, k \in \mathbb{Z} \quad \beta = \alpha + k\pi$$

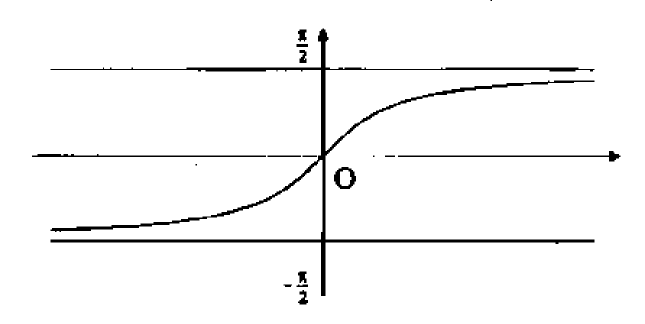


Figure **8.4** – *Fonction arctan*

8-1.5 Exercice : fonctions hyperboliques réciproques

Il ne s'agit pas de fonctions "nouvelles", en ce sens qu'elles s'expriment à l'aide des fonctions "usuelles" (logarithmes, racine carrée etc...). On les rencontre souvent et il est donc utile de connaître leurs définitions :

- **Fonction** argsh **("argument sinus hyperbolique").**

  La fonction impaire sinus hyperbolique :

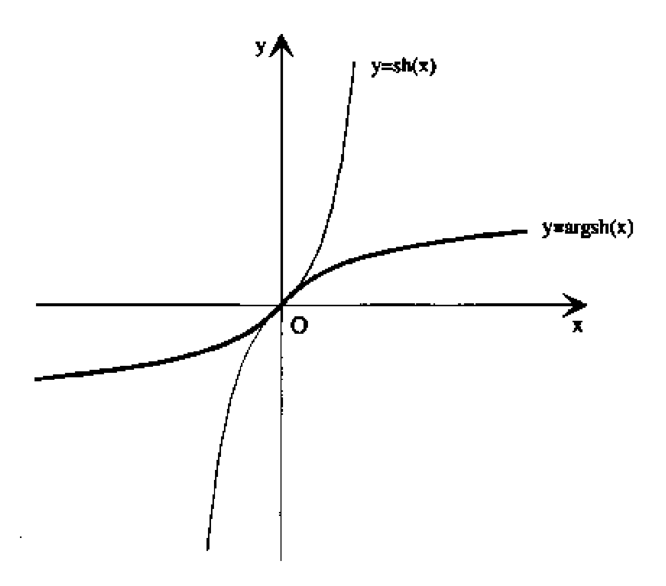


Figure **8.5** – *Fonctions sh et argsh*

$$x \mapsto \operatorname{sh} x = \frac{e^x - e^{-x}}{2}$$

(les anglo-saxons notent plutôt $\sinh x$) est clairement continue $\mathbb{R} \to \mathbb{R}$, strictement croissante et vérifie $\lim_{x \to \pm\infty} \operatorname{sh} x = \pm\infty$. Elle réalise donc un homéomorphisme de $\mathbb{R}$ dans lui-même, et on appelle argsh l'homéomorphisme

réciproque, strictement croissant sur $\mathbb{R}$. Cette fonction s'exprime en fait à l'aide des fonctions usuelles puisque, pour x et $y \in \mathbb{R}$

$$y = \operatorname{sh} x \Leftrightarrow e^{2x} - 2ye^{x} - 1 = 0$$

La quantité e^x est alors solution d'une équation du second degré, dont les racines sont

$$y \pm \sqrt{1+y^2}$$

Comme seule la quantité $y + \sqrt{1+y^2}$ est strictement positive, on obtient l'expression de x en fonction de y :

$$\forall y \in \mathbb{R} \qquad x = \operatorname{argsh} y = \ln\left(y + \sqrt{1+y^2}\right)$$

expression que nous reverrons dans le chapitre consacré aux calculs de primitives.

- **Fonction** argch **("argument cosinus hyperbolique").**

La fonction paire cosinus hyperbolique :

$$x \mapsto \operatorname{ch} x = \frac{e^x + e^{-x}}{2}$$

(les anglo-saxons notent plutôt $\cosh x$) est clairement continue $\mathbb{R} \to \mathbb{R}$ mais n'est plus injective sur $\mathbb{R}$. Elle est strictement croissante sur $\mathbb{R}^+$ (décroissante sur $\mathbb{R}^-$), vérifie $\lim\limits_{x\to+\infty} \operatorname{ch} x = +\infty$ et $\operatorname{ch}(0) = 1$. Elle réalise donc un homéomorphisme de $[0, +\infty[$ dans $[1, +\infty[$, et on appelle argch l'homéomorphisme réciproque, strictement croissant sur $[1, +\infty[$. Cette fonction

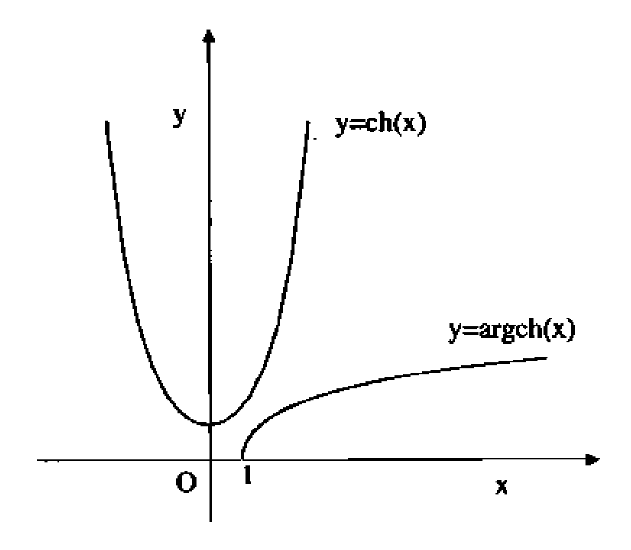


Figure **8.6** *– Fonctions ch et argch*

s'exprime également à l'aide des fonctions usuelles puisque, pour x et $y \in \mathbb{R}$

$$y = \operatorname{ch} x \Leftrightarrow e^{2x} - 2ye^{x} + 1 = 0$$

La quantité e^x est alors solution d'une équation du second degré, dont les racines

$$r = y \pm \sqrt{y^2 - 1}$$

n'existent que si $|y| \geqslant 1$ et sont inverses l'une de l'autre. Comme on veut ensuite résoudre $e^x = r$, il faut de plus avoir $r > 0$, ce qui impose bien $y \in [1, +\infty[$. Comme on cherche enfin une solution $x \geqslant 0$, il faudra prendre $r \geqslant 1$, ce qui impose le signe $+$. En définitive :

$$\forall y \in [1, +\infty[\qquad x = \operatorname{argch} y = \ln\left(y + \sqrt{y^2 - 1}\right)$$

- **Fonction argth (argument tangente hyperbolique).**

La fonction impaire tangente hyperbolique :

$$x \mapsto \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

(les anglo-saxons notent plutôt $\tanh x$) est clairement continue $\mathbb{R} \to \mathbb{R}$, strictement croissante et vérifie $\lim\limits_{x \to \pm\infty} \operatorname{th} x = \pm 1$. Elle réalise donc un homéomorphisme de $\mathbb{R}$ dans $]-1, 1[$, et on appelle argth l'homéomorphisme réciproque, strictement croissant sur $]-1, 1[$.

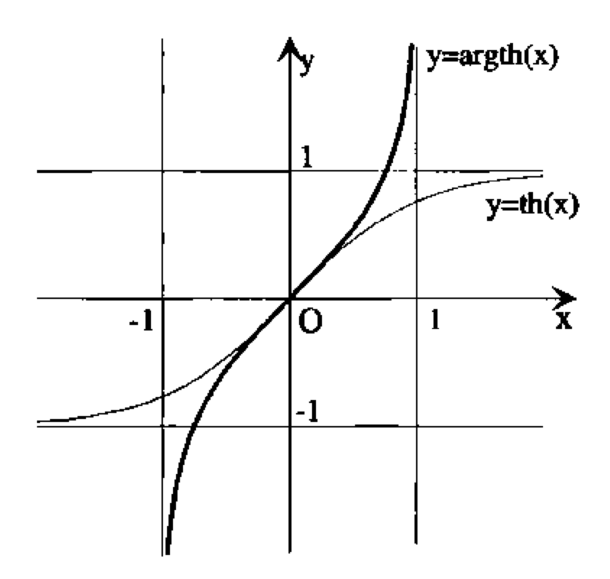


Figure **8.7** – *Fonctions th et argth*

Cette fonction s'exprime en fait à l'aide des fonctions usuelles puisque, pour x et $y \in \mathbb{R}$

$$y = \operatorname{th} x \Leftrightarrow e^{2x}(1 - y) = (1 + y)$$

Pour $y \in]-1, 1[$, on a alors clairement

$$x = \operatorname{argth} y = \frac{1}{2} \ln\left(\frac{1 + y}{1 - y}\right)$$

8-2 Approximation uniforme sur un segment

Nous considérons dans cette section des fonctions $f : [a,b] \to (E, \| \|)$ définies sur un **segment**, à valeurs dans un espace vectoriel normé. Dans le cas de l'approximation de Weierstrass (par des polynômes), on supposera $E = \mathbb{R}$ ou $\mathbb{C}$.

Si $[a,b]$ est un segment, l'espace vectoriel $\mathcal{B}([a,b], E)$ des applications bornées de $[a,b]$ dans E est muni de **la norme de la convergence uniforme**

$$\|f\| = \sup_{x\in[a,b]} \|f(x)\|$$

On sait (voir chapitre sur la continuité) que, si E est complet (ce qui est vrai notamment si $E = \mathbb{R}$ ou $\mathbb{C}$), il en est de même de $(\mathcal{B}([a,b], E), \| \|_\infty)$. On sait de plus que toute fonction continue sur $[a,b]$ y est bornée, et que le sous-espace vectoriel $C^0([a,b], E)$ est **fermé dans** $(\mathcal{B}([a,b], E), \| \|_\infty)$ ("toute limite uniforme de fonctions continues est continue").

DÉFINITION 8-2.1 *Une fonction $s : [a,b] \to E$ est en escalier sur $[a,b]$ si et seulement s'il existe une subdivision de $[a,b]$:*

$$d : \quad a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

telle que

$$\forall i \in \{1, \ldots, n-1\} \qquad s|_{]x_i, x_{i+1}[} \text{ est constante}$$

On dit alors que d est une subdivision adaptée à la fonction s. Il est clair que l'ensemble $\mathcal{E}([a,b], E)$ des fonctions en escalier est un sous-espace vectoriel de $\mathcal{B}([a,b], E)$: la stabilité par multiplication externe est évidente, la stabilité pour la somme est claire si l'on remarque que des subdivisions d et d' adaptées respectivement à deux fonctions en escalier s et s' donnent une subdivision $d \cup d'$ adaptée à la fois à s et s' (donc à $s + s'$...). Si $E = \mathbb{R}$ ou $\mathbb{C}$, $\mathcal{E}([a,b], E)$ est aussi une sous-algèbre de $\mathcal{B}([a,b], E)$.

Remarquons aussi que, si $d : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ est une subdivision adaptée à une fonction en escalier s, on a :

$$s = \sum_{i=0}^{n-1} s\left(\frac{x_i + x_{i+1}}{2}\right) 1_{]x_i, x_{i+1}[} + \sum_{i=0}^{n} s(x_i)\, 1_{\{x_i\}}$$

en notant comme d'habitude 1_A la fonction caractéristique d'une partie A de $[a,b]$ (et en commettant l'abus d'écriture qui consiste à écrire les vecteurs devant les scalaires). En particulier, pour $E = \mathbb{R}$ ou $\mathbb{C}$, $\mathcal{E}([a,b], E)$ est le sous-espace vectoriel de $E^{[a,b]}$ engendré par les fonctions caractéristiques des sous-intervalles de $[a,b]$.

Les fonctions en escalier sont "constantes par intervalles". Nous utiliserons beaucoup ultérieurement les fonctions "continues par intervalles" d'où la définition suivante :

DÉFINITION 8-2.2 *Une fonction $f : [a,b] \to E$ est dite continue par morceaux sur le segment $[a,b]$ si et seulement s'il existe une subdivision de $[a,b]$ (qui sera encore dite adaptée à f)*

$$d: \quad a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

telle que, pour tout $i \in \{1, \dots, n-1\}$, $f|_{]x_i,x_{i+1}[}$ est continue et est prolongeable en une fonction $\mathcal{C}^0$ sur $[x_i, x_{i+1}]$

Cela signifie clairement qu'en chaque point de subdivision, la fonction f possède une limite à droite et une limite à gauche (uniquement à droite en a, à gauche en b) : f ne présente sur $[a,b]$ qu'un nombre fini de discontinuités dites de "première espèce" (c'est-à-dire avec limites à droite et à gauche). On notera $\mathcal{C}^0_m([a,b], E)$ l'ensemble des fonctions continues par morceaux de $[a,b]$ dans E. Il est clair qu'il s'agit encore d'un sous-espace vectoriel de $\mathcal{B}([a,b], E)$, avec

$$\mathcal{E}([a,b], E) \subset \mathcal{C}^0_m([a,b], E) \subset \mathcal{B}([a,b], E)$$

On a évidemment aussi $\mathcal{C}^0([a,b], E) \subset \mathcal{C}^0_m([a,b], E)$. Lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $\mathcal{C}^0_m([a,b], \mathbb{K})$ est clairement une $\mathbb{K}$-algèbre.

8-2.1 Approximation par des fonctions en escalier

THÉORÈME 8-2.3 Si $f : [a,b] \to (E, \| \ \|)$ est continue, on peut l'approcher uniformément sur $[a,b]$ autant qu'on le souhaite par une fonction en escalier :

$$\forall \varepsilon > 0 \quad \exists s \in \mathcal{E}([a,b], E) \qquad \|f - s\|_\infty < \varepsilon$$

Démonstration : Comme f est continue sur le compact $[a,b]$, elle est uniformément continue. Si $\varepsilon > 0$ est donné, il existe $\alpha > 0$ tel que

$$\forall x, y \in [a,b] \qquad |x - y| < \alpha \Rightarrow \|f(x) - f(y)\| < \varepsilon$$

Si on choisit un entier n tel que $\dfrac{b-a}{n} < \alpha$ et une subdivision "à pas constant"

$$d: \ a = x_0 < x_1 = a + \frac{b-a}{n} < \cdots < x_i = a + i\frac{b-a}{n} < \cdots < x_n = b$$

la fonction en escalier

$$s = \sum_{i=0}^{n-1} f(x_i)\, 1_{[x_i,x_{i+1}[} + f(b)\, 1_{\{b\}}$$

répond évidemment à la question. ■

REMARQUE 8-2.4 En prenant $\varepsilon = \dfrac{1}{n}$ pour $n \in \mathbb{N}^*$ et en choisissant s_n dans $\mathcal{E}([a,b], E)$ vérifiant l'inégalité $\|f - s_n\|_\infty < \dfrac{1}{n}$, on construit une **suite de fonctions en escalier qui converge uniformément vers** f sur $[a,b]$. Ce résultat est évidemment équivalent à celui du théorème précédent.

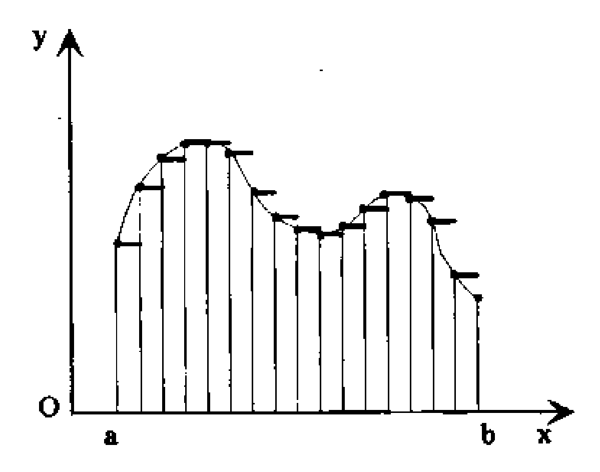


Figure **8.8** – *Approximation uniforme par une fonction en escalier*

COROLLAIRE 8-2.5 Si $f : [a,b] \to (E, \| \|)$ est continue par morceaux, il existe une suite de fonctions en escalier sur $[a,b]$ qui converge uniformément vers f.

Démonstration : Comme précédemment, il suffit de prouver que, pour $\varepsilon > 0$ arbitraire, il existe une fonction s en escalier vérifiant $\|f - s\|_\infty \leqslant \varepsilon$. Si $d : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ est une subdivision adaptée à f, on sait que $f|_{]x_i,x_{i+1}[}$ se prolonge en une fonction continue f_i sur $[x_i, x_{i+1}]$. D'après le théorème qui précède, on peut trouver une fonction $s_i \in \mathcal{E}([x_i, x_{i+1}], E)$ telle que $\|f_i - s_i\|_{\infty,[x_i,x_{i+1}]} \leqslant \varepsilon$. La fonction $s : [a,b] \to E$ définie par

$$\forall i \quad \forall t \in]x_i, x_{i+1}[\quad s(t) = s_i(t) \text{ et } s(t_i) = f(t_i)$$

est clairement en escalier sur $[a,b]$ et vérifie $\|f - s\|_\infty \leqslant \varepsilon$. ∎

On peut caractériser les fonctions $f : [a,b] \to E$ qui sont limites uniformes de suites de fonctions en escalier, lorsque E est un espace complet (donc en particulier lorsqu'on travaille avec des fonctions numériques ou à valeurs dans un espace normé de dimension finie). C'est la notion de fonction réglée développée dans l'exercice qui suit :

EXERCICE 8-2.6 On note $\mathcal{R}([a,b], E)$ l'ensemble des fonctions de $[a,b]$ dans E qui sont limites uniformes de suites de fonctions en escalier. Une fonction $f \in \mathcal{R}([a,b], E)$ est dite **réglée** sur $[a,b]$.

- Montrer que $\mathcal{R}([a,b], E) = \overline{\mathcal{E}([a,b], E)}$, adhérence de $\mathcal{E}([a,b], E)$ dans l'espace vectoriel normé $(\mathcal{B}([a,b], E), \| \|_\infty)$. En déduire que $\mathcal{R}([a,b], E)$ est un sev de $\mathcal{B}([a,b], E)$ (une sous-algèbre dans le cas $E = \mathbb{R}$ ou $\mathbb{C}$).
- Lorsque $(E, \| \|)$ est un espace complet, montrer que toute fonction $f \in \mathcal{R}([a,b], E)$ possède une limite à droite et à gauche en tout point de $[a,b]$. (utiliser le théorème 7-3.10)

- Soit $f : [a,b] \to E$ admettant en tout point de $[a,b]$ une limite à droite et à gauche. Si f n'est pas réglée

$$\exists \varepsilon > 0 \quad \forall s \in \mathcal{E}([a,b],E) \qquad \|f - s\|_\infty < \varepsilon$$

En raisonnant par dichotomie, arriver à une contradiction. Conclure.

- Sans utiliser le résultat précédent, montrer directement qu'une fonction monotone $f : [a,b] \to \mathbb{R}$ est toujours réglée sur $[a,b]$.

8-2.2 Approximation par des fonctions affines par morceaux

Si $f : [a,b] \to E$ est continue, on peut l'approcher par des fonctions en escalier, qui sont discontinues. Si l'on préfère, on peut approcher f par des fonctions "simples", qui ont le bon goût d'être également continues.

DÉFINITION 8-2.7 *$g : [a,b] \to E$ est dite affine par morceaux si et seulement s'il existe une subdivision de $[a,b]$*

$$d : \ a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

telle que, sur chaque intervalle de subdivision $]x_i, x_{i+1}[$, g soit restriction d'une fonction affine $t \mapsto ta_i + b_i$, avec a_i et $b_i \in E$. Si on veut de plus que g soit continue, il suffit d'imposer à g les mêmes conditions sur chacun des segments $[x_i, x_{i+1}]$.

THÉORÈME 8-2.8 Si $f : [a,b] \to E$ est continue, il existe une suite de fonctions continues affines par morceaux sur $[a,b]$ à valeurs dans $\mathbb{E}$ qui converge uniformément vers f.

Démonstration : Comme précédemment, il suffit de trouver, pour $\varepsilon > 0$ arbitraire, une fonction continue affine par morceaux g qui vérifie

$$\|f - g\|_\infty \leqslant \varepsilon$$

Si $\alpha > 0$ est comme dans la démonstration du théorème **8-2**.3 et d la subdivision à pas constant $\frac{b-a}{n} < \alpha$ avec $x_i = a + i\frac{b-a}{n}$, on définira g par intervalles par

$$\forall t \in [0,1] \qquad g(x_i + t(x_{i+1} - x_i)) = (1-t) f(x_i) + t f(x_{i+1})$$

On définit ainsi clairement une fonction continue affine par morceaux, et pour $x \in [x_i, x_{i+1}]$ s'écrivant $x = (1-t)x_i + tx_{i+1}$ avec $t \in [0,1]$, on aura

$$\begin{aligned}\|f(x) - g(x)\| &= \|(1-t)(f(x) - f(x_i)) + t(f(x) - f(x_{i+1}))\| \\ &\leqslant (1-t)\|f(x) - f(x_i)\| + t\|f(x) - f(x_{i+1})\|\end{aligned}$$

ce qui, compte tenu du pas de subdivision, donnera

$$\|f - g\|_\infty \leqslant \varepsilon \quad \blacksquare$$

8-2.3 Approximation polynomiale

Nous énonçons ici, sans démonstration, le théorème de Weierstrass d'approximation uniforme par des polynômes. Nous verrons ultérieurement plusieurs preuves de ce théorème, basées sur les séries de Fourier ou sur la convolution (voir chapitre sur l'intégration). On travaille ici bien évidemment sur des fonctions numériques.

THÉORÈME 8-2.9 Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $f : [a, b] \to \mathbb{K}$ est une fonction continue, il existe une suite de polynômes de $\mathbb{K}[X]$ telles que la suite de fonctions associées converge uniformément vers f sur $[a, b]$.

REMARQUE 8-2.10 On pourrait penser prouver ce théorème en faisant une subdivision de $[a, b]$ à pas constant ($x_i = a + i\dfrac{b-a}{n}$) et considérer le polynôme interpolateur de degré $\leqslant n$ vérifiant

$$\forall i \qquad P_n(x_i) = f(x_i)$$

(*cf.* interpolation de Lagrange). On montre cependant que, même pour des fonctions indéfiniment dérivables sur $[a, b]$, on peut ne pas avoir

$$\lim_{n \to +\infty} \|f - P_n\|_\infty = 0$$

(Ce résultat est connu sous le nom de phénomène de Runge).

8-3 Comparaison de fonctions au voisinage d'un point

Dans tout ce paragraphe, les fonctions sont définies sur une partie A d'un espace normé $(X, \| \|)$, à valeurs dans un espace normé $(E, \| \|)$. On ne fait pas d'hypothèse sur E, mais les applications les plus fréquentes des définitions et théorèmes qui suivent concerneront le cas où E est de dimension finie. On sait alors que toutes les normes sur E sont équivalentes, et que si l'on fixe une base $\mathcal{B} = \{e_1, \ldots, e_p\}$ de E, la convergence d'une suite dans E, l'existence d'une limite d'une fonction à valeurs dans E, etc.. peuvent se lire "coordonnée par coordonnée", en travaillant dans $\mathcal{B}$, ce qui revient en fait à travailler avec la norme dans E

$$\left\| \sum_{i=1}^{p} x_i e_i \right\|_{\infty, \mathcal{B}} = \max_{1 \leqslant i \leqslant p} |x_i|$$

Le cas le plus important est celui des **fonctions numériques**, pour lesquelles $E = \mathbb{R}$ ou $\mathbb{C}$, que nous développerons plus particulièrement dans la suite. On suppose enfin que $a \in X$ est **adhérent** à A, en autorisant le cas a "=" ∞ si A n'est pas bornée dans X, ou $a = +\infty$ (resp. $-\infty$) si A est une partie non majorée (resp. non minorée) de $\mathbb{R}$. Ce dernier cas englobe le cas particulier où $A = \mathbb{N}$, où l'on étudie les **suites** de E.

On étudie alors le comportement des fonctions "au voisinage de a", c'est-à-dire le comportement asymptotique de $f(x)$ pour "x tendant vers a dans A".

8-3.1 Cas des fonctions numériques

Les fonctions sont définies sur $X \supset A$ à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

8-3.1.1 Relation de domination

DÉFINITION 8-3.1 *Si f et g sont deux fonctions $A \to \mathbb{K}$ et $a \in \overline{A}$, on dit que g domine f au voisinage de a ssi*

$$\exists\, V \in \mathcal{V}(a) \quad \exists\, M \in \mathbb{R}^+ \quad \forall\, x \in V \cap A \qquad |f(x)| \leqslant M\,|g(x)| \tag{1}$$

Il s'agit là de la définition la plus générale. Dans la quasi-totalité des cas, lorsque $a \notin A$ et lorsqu'il existe un voisinage V_0 de a tel que g ne s'annule en aucun point de $V_0 \cap A$, la fonction $\frac{f}{g}$ est alors définie sur $V_0 \cap A$, et dire que g domine f en a revient à affirmer[2] que $x \mapsto \frac{f(x)}{g(x)}$ est bornée sur un voisinage de a (ou plus précisément sur sa trace sur A)

Cette définition est la plus intuitive et la plus commode. La définition générale donnée plus haut est donnée pour recouvrir certains cas à la limite de la pathologie, où la fonction g s'annule une infinité de fois sur tout voisinage de a. La propriété (1) montre alors que, au moins sur la trace sur A du voisinage V, les points où g s'annule sont aussi des zéros de la fonction f. On a donc la définition équivalente suivante :

DÉFINITION 8-3.2 *Si f et g sont deux fonctions $A \to \mathbb{K}$ et $a \in \overline{A}$, g domine f au voisinage de a ssi*

$$\exists\, V \in \mathcal{V}(a) \quad \exists\, \Phi \text{ bornée } : V \to \mathbb{C} \quad f = \Phi g \text{ sur } V \cap A$$

Il existe en effet un voisinage V vérifiant la propriété (1). Il suffit alors de définir la fonction Φ par

$$\begin{cases} \Phi(x) = \dfrac{f(x)}{g(x)} \text{ si } x \in V \cap A \text{ et } g(x) \neq 0 \\ \Phi(x) = 1 \text{ si } x \in V \text{ ne vérifie pas cette condition} \end{cases}$$

Notons enfin qu'une relation de domination entre deux fonctions définies sur A est encore valable en restriction à toute partie $B \subset A$ (en supposant évidemment que $a \in \overline{B}$).

Notation : Si g domine f en a, on notera de manière équivalente

$$f \underset{a}{=} O(g) \qquad \text{ou} \qquad f(x) \underset{a}{=} O(g(x)) \qquad \text{ou} \qquad f(x) \underset{x \to a}{=} O(g(x))$$

On omet d'écrire a s'il n'y a pas d'ambiguïté et l'on précise la partie A si nécessaire. On notera que cette "égalité" n'en est pas une, il s'agit en fait d'une appartenance à une classe de fonctions définies sur A et la notation $f \in O(g)$

[2]Si $a \in A$ et $g(a) = 0$, il faudra rajouter la condition $f(a) = 0$

serait plus correcte mais moins utilisable dans des écritures "algébriques" : on écrira par exemple

$$h \underset{a}{=} f + O(g)$$

pour dire

$$h - f \underset{a}{=} O(g)$$

En particulier, l'égalité $O(g) = O(g) + O(g)$ ne signifie pas $1 = 2$ mais que la somme de deux fonctions définies sur A et dominées par g en a est aussi dominée par g. De même, des égalités

$$h = f_1 + O(g) = f_2 + O(g)$$

on ne déduira évidemment pas $f_1 = f_2$, mais $f_1 = f_2 + O(g)$.

EXEMPLE 8-3.3 On a au voisinage de $0 \in \mathbb{R}$

$$\sin(x) - x = O(x^2) \qquad \sin(x) - x = O(x^3) \ \text{ mais } \ \sin(x) - x \neq O(x^4)$$

De même

$$x \sin\left(\frac{1}{x}\right) = O(\sin x) \ \text{ mais } \ x^3 \neq O\left(x \sin\left(\frac{1}{x}\right)\right)$$

Dans ce qui suit, les fonctions sont définies $A \to \mathbb{C}$. Les propriétés suivantes découlent immédiatement des définitions :

- $f = O(1) \Leftrightarrow f$ est bornée au voisinage de a (dans A)
- $f = O(g)$ et $g = O(h) \Rightarrow f = O(h)$ ("transitivité", la réflexivité est évidente)
- $f_1 = O(g)$ et $f_2 = O(g) \Rightarrow f_1 + f_2 = O(g)$ (on domine par la même fonction g)
- $f = O(g) \Rightarrow fh = O(gh)$
- $f_1 = O(g_1)$ et $f_2 = O(g_2) \Rightarrow f_1 f_2 = O(g_1 g_2)$
- Si $\alpha > 0$ est fixé et si $f = O(g)$, alors $f^\alpha = O(g^\alpha)$, sous réserve que les fonctions f^α et g^α soient définies au voisinage de a, ce qui ne pose pas de problème si $\alpha \in \mathbb{N}^*$ mais nécessite que f et g prennent des valeurs réelles positives dans le cas général.

Tout ceci est essentiellement dû au fait que la somme et le produit de deux fonctions bornées au voisinage de a sont encore bornées au voisinage de a.

- Si enfin $f = O(g)$ et s'il existe un voisinage de a sur lequel la fonction f ne s'annule pas, $\frac{1}{f}$ et $\frac{1}{g}$ sont définies au voisinage de a et on a : $\frac{1}{g} = O\left(\frac{1}{f}\right)$ puisque l'égalité (valable au voisinage de a) $f = \Phi g$ avec Φ bornée s'écrit aussi $\frac{1}{g} = \Phi \frac{1}{f}$. On en déduira, avec les mêmes réserves que précédemment, que $g^\alpha = O(f^\alpha)$ si $\alpha < 0$ est un réel fixé.

Terminons par un résultat dans le cas particulier où $A = \mathbb{N}$ et où l'**on compare des suites réelles strictement positives**, qui donne un critère de domination que l'on utilise souvent lorsqu'on étudie des séries.

THÉORÈME 8-3.4 (Comparaison logarithmique) Si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont des suites réelles à termes strictement positifs (éventuellement à partir d'un certain rang) et s'il existe un rang n_0 tel que

$$n \geqslant n_0 \Rightarrow \frac{u_{n+1}}{u_n} \leqslant \frac{v_{n+1}}{v_n}$$

alors on a

$$u_n \underset{+\infty}{=} O(v_n)$$

Démonstration : On peut supposer que u_n et v_n sont > 0 à partir du rang n_0, et on a alors, pour $n > n_0$

$$0 \leqslant u_n = u_{n_0} \prod_{k=n_0}^{n-1} \frac{u_{k+1}}{u_k} \leqslant u_{n_0} \prod_{k=n_0}^{n-1} \frac{v_{k+1}}{v_k} = \frac{u_{n_0}}{v_{n_0}} v_n$$

ce qui donne le résultat (pas de problème dans les manipulations des inégalités puisqu'on travaille avec des réels > 0). ■

Comme la relation de domination (comme dans ce qui suit la relation de négligeabilité) ne fait intervenir que des inégalités entre modules, on a évidemment :

COROLLAIRE 8-3.5 Si $(u_n)_{n\in\mathbb{N}}$ est une suite complexe ne s'annulant pas et si $(v_n)_{n\in\mathbb{N}}$ est réelle à termes strictement positifs (éventuellement à partir d'un certain rang) et s'il existe un rang n_0 tel que

$$n \geqslant n_0 \Rightarrow \left|\frac{u_{n+1}}{u_n}\right| \leqslant \frac{v_{n+1}}{v_n}$$

alors on a

$$u_n \underset{+\infty}{=} O(v_n)$$

EXEMPLE 8-3.6 Le critère de comparaison logarithmique permet de démontrer facilement que, pour n tendant vers $+\infty$, on a

$$\left(\frac{n}{e}\right)^n = O(n!) \quad \text{et} \quad n! = O\left(\left(\frac{n+1}{e}\right)^{n+1}\right)$$

On verra ultérieurement que ces suites sont en fait comparables pour la relation de négligeabilité.

8-3.1.2 Relation de négligeabilité

DÉFINITION 8-3.7 *Si f et g sont deux fonctions $A \to \mathbb{K}$ et $a \in \overline{A}$, on dit que f est négligeable devant g au voisinage de a ssi*

$$\forall\, \eta > 0 \quad \exists\ V \in \mathcal{V}(a) \quad \forall\, x \in V \cap A \qquad |f(x)| \leqslant \eta\, |g(x)| \tag{2}$$

Ici encore, si $a \notin A$ et s'il existe un voisinage V_0 de a tel que g ne s'annule en aucun point de $V_0 \cap A$, la fonction $\dfrac{f}{g}$ est alors définie sur $V_0 \cap A$, et dire que f est négligeable devant g au voisinage de a revient à affirmer[3] que

$$\lim_{\substack{x\to a\\ x\in A}} \frac{f(x)}{g(x)} = 0$$

C'est cette caractérisation qu'on utilise dans 99,99 % des cas. La propriété (2) montre dans le cas général que, si la fonction g s'annule une infinité de fois sur tout voisinage de a, il existe un voisinage de a particulier (par exemple celui correspondant à $\eta = 1$) sur lequel les zéros de g sont aussi des zéros de la fonction f. On a donc la définition équivalente suivante :

DÉFINITION 8-3.8 *Si f et g sont deux fonctions $A \to \mathbb{K}$ et $a \in \overline{A}$, f est négligeable devant g au voisinage de a ssi*

$$\exists\, V_0 \in \mathcal{V}(a) \quad \exists\, \varepsilon : V_0 \to \mathbb{C} \quad \textit{avec } \lim_{x\to a} \varepsilon = 0 \textit{ telle que } \ f = \varepsilon g \textit{ sur } V_0 \cap A$$

La propriété (2) (avec $\eta = 1$) fournit un voisinage V_0 de a sur lequel $|f| \leqslant |g|$, donc où les zéros de g sont aussi zéros de f. Il suffit alors de définir la fonction ε par

$$\begin{cases} \varepsilon(x) = \dfrac{f(x)}{g(x)} \text{ si } x \in V \cap A \text{ et } g(x) \neq 0 \\ \varepsilon(x) = 0 \text{ si } x \in V \text{ ne vérifie pas cette condition} \end{cases}$$

La propriété (2) permet alors de montrer aisément que la fonction ε tend vers 0 en a.

Notation : Si f est négligeable devant g en a, on notera de manière équivalente

$$f \underset{a}{=} o(g) \qquad \text{ou} \qquad f(x) \underset{a}{=} o(g(x)) \qquad \text{ou} \qquad f(x) \underset{x\to a}{=} o(g(x))$$

avec les mêmes remarques que précédemment en ce qui concerne la signification de cette "égalité".

Comme au paragraphe précédent, et essentiellement parce que le produit d'une fonction bornée par une fonction tendant vers 0 tend également vers 0 on a :

- $f \underset{a}{=} o(1) \Leftrightarrow \lim\limits_{\substack{x\to a\\ x\in A}} f(x) = 0$
- $f = o(g) \Rightarrow f = O(g)$ (perte d'information !)

[3]Ici encore, si $a \in A$ et $g(a) = 0$, il faudra rajouter la condition $f(a) = 0$.

- $f = o(g)$ et $g = O(h) \Rightarrow f = o(h)$ (en particulier la négligeabilité est transitive), même conclusion avec $f = O(g)$ et $g = o(h)$
- $f_1 = o(g)$ et $f_2 = o(g) \Rightarrow f_1 + f_2 = o(g)$ (on compare à la même fonction g)
- $f_1 = o(g_1)$ et $f_2 = O(g_2) \Rightarrow f_1 f_2 = o(g_1 g_2)$. En particulier

$$f = o(g) \Rightarrow fh = o(gh)$$

  et on a également $f_1 = o(g_1)$ et $f_2 = o(g_2) \Rightarrow f_1 f_2 = o(g_1 g_2)$
- Si $\alpha > 0$ est fixé et si $f = o(g)$, alors $f^\alpha = o(g^\alpha)$, sous réserve que les fonctions f^α et g^α soient définies au voisinage de a. Si $\alpha < 0$ et si f ne s'annule pas au voisinage de a on aura par contre $g^\alpha = o(f^\alpha)$.

On notera enfin que, si λ est un scalaire non nul, les classes de fonctions (au voisinage de a) $o(g)$ et $o(\lambda g)$ (tout comme les classes $O(g)$ et $O(\lambda g)$) sont égales. Au voisinage de zéro par exemple, pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, les "égalités"

$$f(x) = g(x) + o(x^2) \quad \text{ou} \quad f(x) = g(x) + o(3x^2)$$

sont équivalentes.

On a également le résultat suivant (dit de "changement de variable"), énoncé ici pour la relation de négligeabilité, mais les résultats analogues sont vrais pour la domination et l'équivalence :

PROPOSITION 8-3.9 Si f et $g : A \to \mathbb{K}$ vérifient $f \underset{a}{=} o(g)$, avec $a \in \overline{A}$, si u est une fonction définie sur une partie B d'un evn, avec $u(B) \subset A$, et si $u(x)$ tend vers a lorsque x tend vers $b \in \overline{B}$, alors

$$f(u(x)) \underset{x \to b}{=} o(g(u(x)))$$

(ceci est vrai simplement parce que, si ε est une fonction tendant vers 0 en a, la fonction $\varepsilon \circ u$ tend vers 0 en b).

Par exemple, si f et g sont deux fonctions d'une variable réelle vérifiant

$$f(x) \underset{0}{=} g(x) + o(x^2)$$

on aura alors évidemment

$$f(x \ln x) \underset{0}{=} g(x \ln x) + o(x^2 \ln^2 x)$$

Il serait stupide d'écrire $f(\cos x) \underset{0}{=} g(\cos x) + o(\cos^2 x)$, puisque cette "égalité" décryptée signifierait exactement $\lim\limits_{\substack{x \to 1 \\ x \leqslant 1}} (f - g)(x) = 0$!

A part ces quelques résultats, qui sont presque des évidences pour qui a compris les définitions, **tout autre résultat reliant des opérations entre fonctions et des relations de négligeabilité doit être justifié**, sous peine d'être taxé de nullité (la remarque vaut également pour les relations d'équivalence ou de domination).

Dans l'exemple précédent, on a tenu compte du fait que $\lim\limits_{x \to 0} x \ln x = 0$, conséquence du résultat suivant :

THÉORÈME 8-3.10 (Croissances comparées) Si α et β sont des réels strictement positifs et si k est réel strictement supérieur à 1 on a, pour x réel tendant vers $+\infty$

$$(\ln x)^\alpha = o(x^\beta)$$
$$x^\beta = o(k^x)$$

et, pour n entier tendant vers $+\infty$

$$k^n = o(n!)$$

Preuve : Pour $x \geqslant 1$, on a

$$0 \leqslant \ln x = \int_1^x \frac{dt}{t} \leqslant \int_1^x \frac{dt}{\sqrt{t}} \leqslant 2\sqrt{x}$$

ce qui donne $\ln x \underset{+\infty}{=} O(\sqrt{x})$ et entraîne donc $\ln x = o(x)$. Si $\beta > 0$, on a $\lim\limits_{+\infty} x^\beta = +\infty$ et donc $\ln x^\beta = o(x^\beta)$, soit $\ln(x) = o(x^\beta)$. Enfin, en remplaçant β par $\dfrac{\beta}{\alpha}$ et en élevant à la puissance $\alpha > 0$ on obtient

$$(\ln x)^\alpha \underset{+\infty}{=} o(x^\beta) \tag{1}$$

Si $k > 1$, on a $\lim\limits_{+\infty} k^x = +\infty$, et le résultat précédent donne donc, avec $\beta = 1$, $(\ln k^x)^\alpha = o(k^x)$ soit

$$x^\alpha \underset{+\infty}{=} o(k^x)$$

Enfin, pour comparer les suites k^n et $n!$, on utilise la comparaison logarithmique. Si on choisit $r > k$, on a, à partir d'un certain rang, $r = \dfrac{r^{n+1}}{r^n} \leqslant \dfrac{(n+1)!}{n!}$, ce qui entraîne $r^n = O(n!)$ et donc, puisque $k^n = o(r^n)$ on a

$$k^n = o(n!)$$

Notons enfin que, pour x tendant vers zéro par valeurs strictement positives, $\dfrac{1}{x}$ tend vers $+\infty$ et donc la propriété (1) donne immédiatement

$$(|\ln x|)^\alpha \underset{0}{=} o\left(\frac{1}{x^\beta}\right)$$

donc $\lim\limits_{x \to 0^+} x^\beta (|\ln x|)^\alpha = 0$. ∎

8-3.1.3 Equivalence

DÉFINITION 8-3.11 *Si f et g sont deux fonctions $A \to \mathbb{K}$ et $a \in \overline{A}$, on dit que f est équivalente à g au voisinage de a ssi*

$$f - g \underset{a}{=} o(g)$$

ce qu'on écrit aussi

$$f(x) \underset{a}{\sim} g(x) \Leftrightarrow f(x) \underset{a}{=} g(x) + o(g(x))$$

Comme une fonction négligeable devant g peut s'écrire sous la forme εg sur un voisinage de a, avec $\lim_{a} \varepsilon = 0$, cette définition est équivalente à (en posant $\alpha = 1 + \varepsilon$) :

DÉFINITION 8-3.12 *f est équivalente à g au voisinage de a ssi*

$$\exists\, V_0 \in \mathcal{V}(a) \quad \exists\, \alpha : V_0 \to \mathbb{C} \quad \text{avec } \lim_{x\to a} \alpha = 1 \text{ telle que } f = \alpha g \text{ sur } V_0 \cap A$$

Si la fonction g ne s'annule pas au voisinage de a, ceci équivaut à la caractérisation[4] *(la plus utile en pratique) :*

$$f \underset{a}{\sim} g \Leftrightarrow \lim_{\substack{x\to a\\ x\in A}} \frac{f(x)}{g(x)} = 1$$

Les propriétés suivantes découlent immédiatement des définitions :

- La relation "être équivalente au voisinage de $a \in \overline{A}$" est une relation d'équivalence sur l'ensemble des applications de A dans $\mathbb{K}$. La réflexivité et la transitivité sont évidentes, la symétrie provenant du fait que, si une fonction α tend vers 1 en a, son inverse est définie sur un voisinage de a et tend également vers 1.

- Si $f \underset{a}{\sim} g$ et si g possède une limite $l \in \mathbb{K}$ en a, f tend aussi vers l en a. C'est une conséquence évidente des théorèmes sur les limites. Ce résultat est encore valable avec $l = \pm\infty$ pour des fonctions à valeurs réelles ou avec $l = \infty$ pour des fonctions complexes. Pour "calculer" la limite (c'est-à-dire, pour être précis, prouver l'existence d'une limite et trouver sa valeur) de f en a, on peut remplacer f par une fonction équivalente g et résoudre le problème pour g.

- Si $l \in \mathbb{K} - \{0\}$, on a $f \underset{a}{\sim} l \Leftrightarrow \lim_{x\to a} f(x) = l$

Par contre $f \underset{a}{\sim} 0 \Leftrightarrow f$ est identiquement nulle au voisinage de a

  un raisonnement invoquant explicitement (ou implicitement) une équivalence avec la fonction nulle contient (presque) toujours une mauvaise utilisation du symbole $\sim$ et est suspect.

- La relation d'équivalence est compatible avec les produits et les quotients :

$$\text{si } f_1 \underset{a}{\sim} f_2 \text{ et } g_1 \underset{a}{\sim} g_2 \text{ alors } f_1 g_1 \underset{a}{\sim} f_2 g_2 \text{ et } \frac{f_1}{g_1} \underset{a}{\sim} \frac{f_2}{g_2}$$

  sous réserve (pour le quotient) que la fonction g_1 ne s'annule pas sur un voisinage de a (sauf peut-être en a) ce qui assure qu'il en est de même pour la fonction g_2. De même, si α est un réel fixé

$$f_1 \underset{a}{\sim} f_2 \Rightarrow f_1^\alpha \underset{a}{\sim} f_2^\alpha \text{ et } |f_1| \underset{a}{\sim} |f_2|$$

  sous réserve de définition des fonctions.

[4]Si $g(a) = 0$, on doit avoir $f(a) = 0$ et la limite est alors prise suivant $A - \{a\}$.

• Comme pour la relation de domination on a

PROPOSITION 8-3.13 Si f et $g : A \to \mathbb{K}$ vérifient $f \underset{a}{\sim} g$, avec $a \in \overline{A}$, si u est une fonction définie sur une partie B d'un evn, avec $u(B) \subset A$, et si $u(x)$ tend vers a lorsque x tend vers $b \in \overline{B}$, alors

$$f(u(x)) \underset{x \to b}{\sim} g(u(x))$$

(puisque si α est une fonction tendant vers 1 en a, la fonction $\alpha \circ u$ tend vers 1 en b).

Par contre il faut se garder d'additionner (ou de faire des combinaisons linéaires d'équivalence). En particulier

$$\boxed{f_1 \underset{a}{\sim} f_2 \text{ et } g_1 \underset{a}{\sim} g_2 \not\Rightarrow f_1 + g_1 \underset{a}{\sim} f_2 + g_2}$$

Par exemple, s'il est correct (mais maladroit, et cela prouve qu'on a mal compris la définition de l'équivalence) d'écrire $\cos x \underset{0}{\sim} 1 - \dfrac{x^2}{2}$, en déduire que

$$\cos x - 1 \underset{0}{\sim} -\frac{x^2}{2}$$

est une grave erreur de raisonnement. Ecrire $\cos x \underset{0}{\sim} 1 - \dfrac{x^2}{2}$, c'est en effet ni plus ni moins qu'écrire $\cos x \underset{0}{\sim} 1$ ou encore $\cos x \underset{0}{\sim} 1 + \dfrac{x^5}{27}$ (transitivité de l'équivalence). Espérer déduire de cette écriture un équivalent de $\cos x - 1$ est illusoire (et extrêmement dangereux) !

On a cependant le résultat suivant, lorsqu'on compare deux fonctions à une même troisième :

PROPOSITION 8-3.14 Si f_1, f_2 et $g : A \to \mathbb{K}$ vérifient

$$f_1 \underset{a}{\sim} \lambda_1 g \quad \text{et} \quad f_2 \underset{a}{\sim} \lambda_2 g$$

où λ_1 et λ_2 sont deux constantes (non nulles !) alors

$$\lambda_1 + \lambda_2 \neq 0 \Rightarrow f_1 + f_2 \underset{a}{\sim} (\lambda_1 + \lambda_2) g$$

Démonstration : Il suffit de traduire les équivalences par

$$f_1(x) \underset{a}{=} \lambda_1 g(x) + o(g(x)) \quad \text{et} \quad f_2(x) \underset{a}{=} \lambda_2 g(x) + o(g(x))$$

pour en déduire

$$f_1(x) + f_2(x) \underset{a}{=} (\lambda_1 + \lambda_2) g(x) + o(g(x))$$

ce qui peut se traduire, **si la somme $\lambda_1 + \lambda_2$ est non nulle**, par :

$$f_1 + f_2 \underset{a}{\sim} (\lambda_1 + \lambda_2) g$$

Si $\lambda_1 + \lambda_2 = 0$ (on dit parfois que "les parties principales se détruisent"), la conclusion est alors $f_1 + f_2 \underset{a}{=} o(g)$. Un conseil dans cette situation, utiliser les symboles d'égalité (sans oublier les $o(g)$!) plutôt que les $\sim$, puisque cela justifie entièrement le résultat et a l'avantage de donner un résultat correct même lorsque les parties principales se détruisent. ■

Attention aussi à l'erreur fréquente :

$$\boxed{f(x) \underset{x\to a}{\sim} g(x) \not\Leftrightarrow f(x) - g(x) \underset{x\to a}{\to} 0 \quad \text{(en général)}}$$

S'il y avait équivalence, cela signifierait qu'au voisinage de a les classes de fonctions $o(1)$ et $o(g)$ sont égales !

Toute composition d'équivalences à gauche par une même fonction (autre que $x \mapsto x^\alpha$ avec α fixe) est à proscrire sans étude particulière. Par exemple, pour étudier l'équivalence de deux fonctions de la forme $e^{f(x)}$ et $e^{g(x)}$, on reviendra à la définition, en étudiant le quotient $e^{f(x)-g(x)}$. De même, pour étudier

$$\lim_{n\to+\infty} u_n \text{ avec } u_n = n\ln\left(1 + \frac{n}{n^2-1}\right)$$

on n'écrira pas $\dfrac{n}{n^2-1} \sim \dfrac{1}{n}$ donc $\ln\left(1 + \dfrac{n}{n^2-1}\right) \sim \ln(1+\dfrac{1}{n}) \sim \dfrac{1}{n}$ donc $u_n \sim n\dfrac{1}{n}$ et $\lim u_n = 1$ (où est la faute ?) mais $\ln(1+u) \underset{0}{\sim} u$ et $\lim\limits_{n\to+\infty} \dfrac{n}{n^2-1} = 0$ donc $\ln\left(1 + \dfrac{n}{n^2-1}\right) \sim \dfrac{n}{n^2-1} \sim \dfrac{1}{n}$ et $\lim u_n = 1$.

Dans le cas de fonctions réelles, la notion de fonctions équivalentes permet de contrôler souvent le signe d'une fonction au voisinage d'un point (sans avoir de renseignement précis sur la "taille" du voisinage où l'étude du signe est valable) :

THÉORÈME 8-3.15 Si f et $g : A \to \mathbb{R}$ vérifient $f \underset{a}{\sim} g$ avec $a \in \overline{A}$, il existe un voisinage V de a tel que pour tout $x \in V \cap A$ on ait

$$f(x) = 0 \Leftrightarrow g(x) = 0 \quad \textbf{et} \quad f(x) \neq 0 \Rightarrow f(x) \textbf{ et } g(x) \textbf{ ont même signe}$$

Démonstration : $\exists\ V_0 \in \mathcal{V}(a)\quad \exists\ \alpha : V_0 \to \mathbb{R}$ avec $\lim_{x\to a} \alpha = 1$ telle que $f = \alpha g$ sur $V_0 \cap A$. Comme α tend vers 1 en a, on peut trouver un voisinage V de a vérifiant $V \subset V_0$ tel que $\alpha > 0$ sur $V \cap A$. Le théorème en découle. ■

On remarquera enfin que si $f \underset{a}{\sim} g$, les classes de fonctions (pour l'étude au voisinage de a) $o(f)$ et $o(g)$ sont égales. Il en est de même des classes $O(f)$ et $O(g)$.

8-3.2 Généralisation aux fonctions vectorielles

8-3.2.1 Domination, négligeabilité

On compare ici une fonction vectorielle $f : A \to (E, \| \ \|)$ (où E est un $\mathbb{K}$-ev) à une fonction numérique $\varphi : A \to \mathbb{K}$. Ce qui importe en fait, c'est la fonction $x \mapsto \|f(x)\|$, qu'on note simplement $\|f\|$ (s'il n'y a pas d'ambiguïté), que l'on compare à la fonction φ (c'est-à-dire à la fonction $|\varphi|$).

DÉFINITION 8-3.16 *On dit que f est **dominée par** φ au voisinage de a ssi*

$$\|f\| \underset{a}{=} O(\varphi)$$

c'est-à-dire s'il existe une majoration de la forme

$$\|f(x)\| \leqslant M\,|\varphi(x)|$$

(avec M constante), valable sur un voisinage de a. Comme dans le paragraphe précédent, il est facile de voir que ceci équivaut à une égalité de la forme

$$f(x) = \varphi(x)g(x)$$

valable sur un voisinage de a, où g est une fonction vectorielle (à valeurs dans le même espace E que f) bornée au voisinage de a.

DÉFINITION 8-3.17 *On dit que f est **négligeable devant** φ au voisinage de a ssi*

$$\|f\| \underset{a}{=} o(\varphi)$$

c'est-à-dire si on a la propriété

$$\forall\, \eta > 0 \quad \exists\, V \in \mathcal{V}(a) \quad \forall\, x \in V \cap A \quad \|f(x)\| \leqslant \eta\,|\varphi(x)|$$

ce qui équivaut ici à une égalité de la forme

$$f(x) = \varphi(x)\varepsilon(x)$$

valable sur un voisinage de a, où ε est une fonction vectorielle (à valeurs dans le même espace E que f) tendant vers 0_E en a.

Si la fonction φ ne s'annule pas au voisinage de a (sauf peut-être en a) cela signifie que $\dfrac{1}{\varphi} f$ tend vers 0_E en a dans le cas de la négligeabilité, est bornée au voisinage de a pour la domination. Si on garde à l'esprit que l'on compare des fonctions vectorielles à des fonctions numériques, les principales propriétés énoncées dans le cas des fonctions numériques s'étendent à ce nouveau contexte (exercices). Aucun problème pour les résultats relatifs aux sommes et combinaisons linéaires (qui sont des opérations d'espaces vectoriels). On notera que la notion de produit est ici remplacée par la notion d'application bilinéaire continue :

THÉORÈME 8-3.18 Si $f : A \to E_1$ et $g : A \to E_2$ sont à valeurs dans deux espaces normés, si $B : E_1 \times E_2 \to F$ est une application bilinéaire continue et si (par exemple) $f(x) \underset{a}{=} o(\varphi(x))$ et $g(x) \underset{a}{=} O(\psi(x))$ où φ et ψ sont des fonctions numériques $A \to \mathbb{K}$, alors $B(f(x), g(x)) \underset{a}{=} o(\varphi\psi(x))$.

Démonstration : Il existe une constante $k > 0$ telle que

$$\forall (x_1, x_2) \in E_1 \times E_2 \quad \|B(x_1, x_2)\| \leqslant k \|x_1\| \|x_2\|$$

ce qui traduit la continuité de B. On en déduit

$$\|B(f(x), g(x))\| \underset{a}{=} O(\|f(x)\| \|g(x)\|)$$

d'où immédiatement le résultat puisque

$$\|f(x)\| \|g(x)\| \underset{a}{=} o(\varphi(x)\psi(x)) \quad \blacksquare$$

REMARQUE 8-3.19 Lorsque l'espace d'arrivée E est de dimension finie p, en fixant une base $\mathcal{B} = \{e_1, \dots, e_p\}$, l'application $f : A \to E$ est définie par les applications coordonnées $(f_i)_{1 \leqslant i \leqslant p} : A \to \mathbb{K}$ telles que $f(x) = \sum f_i(x) e_i$. Comme rappelé en tête de cette section, les normes sur E étant toutes équivalentes[5] , on a

$$\forall i \quad |f_i| = O(\|f\|) \text{ et } \|f\| = O\left(\sum_{i=1}^{p} |f_i|\right)$$

Il en découle facilement que

$$\begin{array}{lcl} f = O(\varphi) & \Leftrightarrow & \forall i \quad f_i = O(\varphi) \\ f = o(\varphi) & \Leftrightarrow & \forall i \quad f_i = o(\varphi) \end{array}$$

8-3.2.2 Equivalence

On compare ici des fonctions à valeurs dans le même espace vectoriel normé.

DÉFINITION 8-3.20 *Si f et $g : A \to E$ sont deux fonctions vectorielles, on dit que f est équivalente à g au voisinage de $a \in \overline{A}$ ssi*

$$f(x) - g(x) \underset{a}{=} o(\|g(x)\|)$$

Il existe donc une fonction vectorielle ε définie au voisinage de a (dans A) tendant vers 0_E en a telle que

$$f(x) = g(x) + \|g(x)\| \varepsilon(x)$$

Cette relation est évidemment réflexive. Comme

$$\|f(x)\| - \|g(x)\| = O(\|f(x) - g(x)\|)$$

[5] Ne pas se tromper ici sur le sens du mot "équivalentes". Il s'agit ici de comparer des normes. Si deux normes sont équivalentes, elles sont "O" l'une de l'autre au voisinage de 0_E.

lorsque $f \underset{a}{\sim} g$, on aura $\|f(x)\| - \|g(x)\| \underset{a}{=} o(\|g(x)\|)$ et donc $\|f(x)\| \underset{a}{\sim} \|g(x)\|$. On en déduit que la relation est symétrique et transitive puisque

$$f - g = o(\|g\|) \text{ et } g - h = o(\|g\|) \Rightarrow f - h = o(\|g\|) = o(\|h\|) \Rightarrow f \sim h$$

La relation d'équivalence n'est pas compatible avec les opérations de l'espace vectoriel E (combinaisons linéaires) et n'est en conséquence pas très utilisable sans précaution. La propriété relative au produit se généralise :

THÉORÈME 8-3.21 Si $f_1, g_1 : A \to E_1$ et $f_2, g_2 : A \to E_2$ sont à valeurs dans deux espaces normés, si $B : E_1 \times E_2 \to F$ est une application bilinéaire continue et si $f_1(x) \underset{a}{\sim} g_1(x)$ et $f_2(x) \underset{a}{\sim} g_2(x)$ alors $B(f_1(x), f_2(x)) \underset{a}{\sim} B(g_1(x), g_2(x))$.

Démonstration : Exercice.

On notera, pour terminer qu'il n'est pas question (sans hypothèses supplémentaires) de donner un sens au rapport $\dfrac{f}{g}$ puisque les vecteurs $f(x)$ et $g(x)$ ne sont pas colinéaires en général. L'équivalence ne signifie donc pas l'existence d'une fonction numérique α tendant vers 1 en a telle qu'on ait l'égalité $f = \alpha g$ au voisinage de a.

Attention ! Comme la relation "être équivalent à" n'est pas compatible avec les opérations de l'espace vectoriel E, il n'y a pas de propriété analogue à celle énoncée à la remarque (**8-3**.19). En particulier, il faut se garder de croire que l'équivalence puisse se lire "coordonnée par coordonnée", comme le montre l'exemple, au voisinage de 0, avec $E = \mathbf{R}^2$

$$f(t) = (t, t^2) \sim g(t) = (t + t^2, t^3)$$

8-4 Dérivabilité

8-4.1 Définition. Propriétés élémentaires

DÉFINITION 8-4.1 *Soit f définie sur un voisinage de $a \in \mathbf{R}$, à valeurs dans un espace vectoriel normé $(E, \|\ \|)$. On dit que f est dérivable en a ssi*

$$\lim_{\substack{t \to a \\ t \neq a}} \frac{f(t) - f(a)}{t - a} = l \text{ existe dans } (E, \|\ \|)$$

On dit alors que l est le vecteur dérivé de f en a, et on note $l = f'(a)$. Il est clair que f est dérivable en a, de vecteur dérivé égal à l si et seulement si

$$f(t) \underset{a}{=} f(a) + (t - a)l + o(t - a) \quad \text{ou} \quad f(a + h) \underset{0}{=} f(a) + hf'(a) + o(h)$$

L'application affine $T_a : \mathbb{R} \to E$ définie par

$$T_a(t) = f(a) + (t - a)l$$

est tangente[6] à f en a. L'application linéaire de $\mathbb{R}$ vers E associée

$$df_a : t \mapsto tl = tf'(a)$$

est appelée **différentielle** de f en a.

L'interprétation géométrique du vecteur dérivé est classique : si $t \mapsto f(t)$ est une fonction continue au voisinage de a, le point $f(t)$ décrit le support d'un arc paramétré dans l'espace normé E. Pour $t \neq a$ et en supposant $f(t) \neq f(a)$, le vecteur $\dfrac{f(t) - f(a)}{t - a}$ est un vecteur directeur de la droite "sécante" $\mathcal{S}_t$ passant par les points $f(t)$ et $f(a)$. Lorsque $f'(a)$ est non nul, on peut considérer que la droite affine

$$\mathcal{D}_a = f(a) + \operatorname{vect} f'(a)$$

est position limite de $\mathcal{S}_t$ pour $t \to a$. On dit que $\mathcal{D}_a$ est la **tangente** à l'arc paramétré $t \mapsto f(t)$ au point de paramètre a.

Lorsque $E = \mathbb{R}$, le rapport

$$\frac{f(t) - f(a)}{t - a}$$

est aussi le coefficient directeur de la droite $M_a M_t$ où M_t est le point de coordonnées $(t, f(t))$ décrivant la représentation graphique de f dans le plan rapporté à un repère affine $(O, \vec{\imath}, \vec{\jmath})$. Le nombre dérivé $f'(a)$ est alors coefficient directeur de la tangente en M_a à cette représentation graphique. L'équation de cette tangente est donc

$$(\mathcal{D}_a) \qquad Y - f(a) = f'(a)(X - a)$$

Si f est définie sur un voisinage de a à droite, et si le taux d'accroissement $\dfrac{f(t) - f(a)}{t - a}$ possède une limite à droite en a, on dira que f est dérivable à droite en a, et cette limite sera notée $f'_d(a)$. Définition identique à gauche.

PROPOSITION 8-4.2 Si f est dérivable en a, elle est continue en a. La réciproque est fausse comme le montre le contre exemple

$$f(t) = t \sin \frac{1}{t}, \textbf{ avec } f(0) = 0, \textbf{ continue non dérivable en } a = 0$$

PROPOSITION 8-4.3 Si f est à valeurs dans un espace normé de dimension finie n, rapporté à une base $B = (e_i)_{1 \leqslant i \leqslant n}$, en faisant intervenir les fonctions coordonnées f_i, on peut écrire

$$f = \sum_{i=1}^{n} f_i e_i$$

[6]Deux applications f et g définies au voisinage de $a \in \mathbb{R}$ à valeurs dans E sont dites tangentes en a ssi

$$f(t) - g(t) \underset{a}{=} o(t - a)$$

Il est clair (exercice) qu'il peut exister au plus une application affine tangente à une fonction f en a, et que supposer cette existence revient à supposer la dérivabilité de f en a.

La fonction f est alors dérivable en a ssi les f_i le sont, et on a alors

$$f'(a) = \sum_{i=1}^{n} f_i'(a)\, e_i$$

En particulier, une fonction complexe est dérivable ssi sa partie réelle et sa partie imaginaire le sont.

Si f est dérivable en tout point d'une partie X de son domaine de définition, on peut définir sur X la **fonction dérivée**

$$X \ni t \mapsto f'(t) \in E$$

Remarquons à ce propos que faire l'hypothèse de l'existence de $f'(a)$, c'est déjà supposer que f est définie au **voisinage** de a (éventuellement à droite ou à gauche par abus de langage).

8-4.2 Calcul des dérivées

8-4.2.1 Linéarité

THÉORÈME 8-4.4 Si $f, g : I \to E$ sont définies au voisinage de a et sont dérivables en a, il en est de même de toute combinaison linéaire de f et g, et

$$\forall\, \lambda, \mu \in \mathbb{K} \qquad (\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$$

THÉORÈME 8-4.5 Soient $(E, \|\ \|)$ et $(F, \|\ \|)$ deux $\mathbb{K}$-ev normés et $\Phi : E \to F$ une application linéaire continue. Si $f : I \to E$ est définie au voisinage de a et dérivable en a, il en est de même de $\Phi \circ f : I \to F$, et on a

$$(\Phi \circ f)'(a) = \Phi(f'(a))$$

Démonstration : On a, pour $h \in \mathbb{R}$ voisin de 0

$$f(a+h) = f(a) + hf'(a) + h\varepsilon(h) \quad \text{avec} \quad \lim_{h \to 0} \varepsilon(h) = 0_E$$

Comme Φ est linéaire, on a alors

$$\Phi \circ f(a+h) = \Phi \circ f(a) + h\Phi(f'(a)) + h\Phi(\varepsilon(h))$$

Comme Φ est continue, on a $\lim_{h \to 0} \Phi(\varepsilon(h)) = 0_F$, ce qui prouve le résultat. ■

8-4.2.2 Composition

Nous composerons ici une fonction numérique à valeurs réelles avec une fonction vectorielle :

THÉORÈME 8-4.6 Soit $\varphi : I \to \mathbb{R}$ définie au voisinage de a, dérivable en a et $f : J \to E$ définie sur un voisinage de $\varphi(a)$, dérivable en $\varphi(a)$. La fonction $f \circ \varphi$ est alors définie au voisinage de a, est dérivable en a et

$$(f \circ \varphi)'(a) = \varphi'(a) . f'(\varphi(a))$$

Démonstration : Comme φ est dérivable en a, elle est continue en a. Si $\eta > 0$ est tel que $]\varphi(a) - \eta, \varphi(a) + \eta[\subset J$, on peut trouver $\alpha > 0$ suffisamment petit pour avoir

$$]a - \alpha, a + \alpha[\subset I \text{ et } |t - a| < \alpha \Rightarrow |\varphi(t) - \varphi(a)| < \eta$$

ce qui montre que $f \circ \varphi$ est définie au voisinage de a.

Pour $|h| < \alpha$, on peut écrire

$$\varphi(a + h) = \varphi(a) + h\varphi'(a) + h\varepsilon(h) \quad \text{avec} \quad \lim_{h \to 0} \varepsilon(h) = 0$$

et de même, pour $|k| < \eta$

$$f(\varphi(a) + k) = f(\varphi(a)) + kf'(\varphi(a)) + k\varepsilon_1(k) \text{ avec } \lim_{k \to 0} \varepsilon_1(k) = 0_E$$

Comme par hypothèse, pour $|h| < \alpha$, $|\varphi(a + h) - \varphi(a)| < \eta$, on peut alors écrire

$$f(\varphi(a + h)) = f(\varphi(a)) + (h\varphi'(a) + h\varepsilon(h)) f'(\varphi(a)) + (h\varphi'(a) + h\varepsilon(h)) \varepsilon_1(h\varphi'(a) + h\varepsilon(h))$$

ce qui peut clairement s'écrire

$$f(\varphi(a + h)) = f(\varphi(a)) + h\varphi'(a) f'(\varphi(a)) + h\varepsilon_2(h)$$

avec $\lim_{h \to 0} \varepsilon_2(h) = 0_E$, ce qui prouve le résultat. ■

REMARQUE 8-4.7 Avec la notation des différentielles, le théorème précédent peut s'écrire

$$d(f \circ \varphi)_a = df_{\varphi(a)} \circ d\varphi_a$$

puisque le second membre de cette égalité est l'application linéaire $\mathbb{R} \to E$ qui au réel h associe le vecteur $h\varphi'(a) f'(\varphi(a))$. C'est cette formule qui se généralisera lorsqu'on composera entre elles des fonctions vectorielles différentiables (voir le chapitre de calcul différentiel à plusieurs variables).

8-4.2.3 Effet d'une application p-linéaire continue

THÉORÈME 8-4.8 Soient $E_1, \ldots, E_p$ et F des espaces vectoriels normés, et

$$\Phi : \prod_{i=1}^{p} E_i \to F$$

une application p-linéaire continue. Si, pour $1 \leqslant i \leqslant p$

$$f_i : I \to E_i$$

est une application dérivable en a, la fonction

$$\Psi : I \to F \textbf{ définie par } t \mapsto \Phi(f_1(t), \ldots, f_p(t))$$

est dérivable en a et

$$\Psi(a) = \sum_{i=1}^{p} \Phi(f_1(a), \ldots, f_{i-1}(a), f_i'(a), f_{i+1}(a), \ldots, f_p(a))$$

Démonstration : Rappelons que la continuité de Φ se traduit par l'existence d'une constante $C > 0$ telle que

$$\forall (y_1, \ldots, y_p) \in \prod_{i=1}^{p} E_i \qquad \|\Phi(y_1, \ldots, y_p)\| \leqslant C \prod_{i=1}^{p} \|y_i\|$$

Cette condition est forcément vérifiée si les p espaces E_i sont tous de dimension finie. Ecrivons la dérivabilité des fonctions f_i :

Pour $|h| < \alpha$, on a

$$f_i(a+h) = f_i(a) + hf_i'(a) + h\varepsilon_i(h) \quad \text{avec} \quad \lim_{h\to 0} \varepsilon_i(h) = 0_{E_i}$$

On écrit ainsi $f_i(a+h)$ comme somme de trois termes. Lorsqu'on développe par p-linéarité $\Psi(a+h) = \Phi(f_1(a+h), \ldots, f_p(a+h))$, on obtient 3^p termes. Le premier est obtenu en prenant $f_i(a)$ dans chaque argument de Φ, il vaut donc $\Psi(a)$. Viennent ensuite p termes obtenus en prenant dans un argument de Φ le terme $hf_i'(a)$ et dans les autres arguments les termes en $f_j(a)$. La somme de ces p termes vaut

$$h\left[\sum_{i=1}^{p} \Phi(f_1(a), \ldots, f_{i-1}(a), f_i'(a), f_{i+1}(a), \ldots, f_p(a))\right]$$

Il est aisé de voir que chacun des termes restants est un $O(h^2)$, donc aussi un $o(h)$, puisque chacun de ces termes s'écrit

$$\Delta(h) = \Phi(y_1(h), \ldots, y_p(h))$$

avec

$$\|\Delta(h)\| \leqslant C \prod_{i=1}^{p} \|y_i(h)\|$$

les fonctions $h \mapsto y_i(h)$ étant toutes bornées au voisinage de 0, deux (au moins) d'entre elles étant des $O(h)$. On a donc bien

$$\Psi(a+h) = \Psi(a) + h\left[\sum_{i=1}^{p} \Phi(f_1(a), \ldots, f_{i-1}(a), f_i'(a), f_{i+1}(a), \ldots, f_p(a))\right] + o(h)$$

et le résultat annoncé est ainsi prouvé. ■

Les applications de ce résultat sont multiples :

- Dérivée d'un produit de fonctions numériques (dans ce cas, la fonction Φ est alors définie sur $\mathbb{K}^p$ par $(x_1, \ldots, x_p) \mapsto \prod_{i=1}^{p} x_i$)

$$\left(\prod_{i=1}^{p} u_i\right)'(a) = \sum_{i=1}^{p} u_1(a) \cdots u_{i-1}(a)\, u_i'(a)\, u_{i+1}(a) \cdots u_p(a)$$

  En particulier $(u^p)'(a) = pu^{p-1}(a)\, u'(a)$.

- Dérivée d'un déterminant. Rappelons que c'est une technique parfois utilisée pour calculer des déterminants dépendant d'un paramètre.
- Dérivée d'un produit scalaire ($\mathbb{E}$ est un espace euclidien et $\Phi : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$, $(x,y) \mapsto \langle x,y \rangle$)

$$\langle u,v \rangle' (a) = \langle u',v \rangle (a) + \langle u,v' \rangle (a)$$

- Dérivée d'un produit vectoriel dans un espace vectoriel euclidien orienté de dimension 3

$$(u \wedge v)' (a) = (u' \wedge v) (a) + (u \wedge v') (a)$$

EXERCICE 8-4.9 Si $A \in \mathcal{M}_n(\mathbb{C})$, déterminer le coefficient du terme en X de $\chi_A(X)$. Indication : c'est $\chi'_A(0)$...

8-4.2.4 Inverse et quotient (fonctions numériques)

THÉORÈME 8-4.10 Si $f : I \to \mathbb{K}$ est définie au voisinage de a avec $f(a) \neq 0$, et est dérivable en a, alors $\frac{1}{f}$ est définie au voisinage de a et est dérivable en a avec

$$\left(\frac{1}{f}\right)' (a) = -\frac{f'(a)}{f^2(a)}$$

Démonstration : Comme $f(a) \neq 0$, la dérivabilité de f en a entraînant sa continuité en a, la fonction f ne s'annule pas sur un voisinage de a, ce qui permet d'y définir la fonction $\frac{1}{f}$. On a alors, pour $|h|$ suffisamment petit

$$\lim_{h \to 0} \frac{1}{h}\left(\frac{1}{f(a+h)} - \frac{1}{f(a)}\right) = \lim_{h \to 0} -\left(\frac{f(a+h)-f(a)}{h}\right) . \frac{1}{f(a+h)\, f(a)} = -\frac{f'(a)}{f^2(a)} \quad \blacksquare$$

COROLLAIRE 8-4.11 Avec les hypothèses précédentes, si $g : I \to \mathbb{K}$ est dérivable en a, il en est de même de $\frac{g}{f}$, et on a

$$\left(\frac{g}{f}\right)' (a) = \frac{f(a)\, g'(a) - g(a)\, f'(a)}{f^2(a)}$$

Il s'agit simplement ici de la formule donnant la dérivée du produit $g \times \frac{1}{f}$, après réduction au même dénominateur. Remarquons que, souvent, il est préférable de dériver ce quotient comme un produit

$$\left(\frac{g}{f}\right)' = g' \left(\frac{1}{f}\right) + g \left(\frac{1}{f}\right)'$$

et de réduire ensuite au même dénominateur, si nécessaire. Par exemple, le dénominateur "raisonnable" pour la dérivée de $\frac{g}{f^n}$ est f^{n+1} et pas f^{2n}.

8-4.2.5 Dérivée d'une réciproque (fonctions réelles strictement monotones)

THÉORÈME 8-4.12 Soit I, J deux intervalles réels et $f : I \to J$ un homéomorphisme d'intervalles. On suppose que f est dérivable en un point $a \in I$. Pour que l'homéomorphisme réciproque f^{-1} soit dérivable en $f(a)$, il faut et il suffit que $f'(a) \neq 0$. On a alors

$$\left(f^{-1}\right)'(f(a)) = \frac{1}{f'(a)}$$

Démonstration : Le résultat est prévisible, puisque les représentations graphiques de f et f^{-1} sont, en repère orthonormé, symétriques par rapport à la première bissectrice des axes, et parce que deux droites symétriques par rapport à cette bissectrice ont des pentes inverses.

La condition $f'(a) \neq 0$ est nécessaire pour avoir la dérivabilité de $g = f^{-1}$ en $b = f(a)$, puisque $g \circ f = id_I$ et, si l'on suppose g dérivable en b, on a, d'après le résultat de composition,

$$1 = (g \circ f)'(a) = g'(b)\, f'(a)$$

Ceci entraîne bien $f'(a) \neq 0$ et montre que $g'(b)$ et $f'(a)$ sont inverses.

Réciproquement, supposons $f'(a) \neq 0$ et envisageons le cas (par exemple) où f est strictement croissante. Supposons enfin que a soit un point intérieur à I, ce qui entraîne que $b = f(a)$ est intérieur à J. Soit $\alpha > 0$ avec $[a - \alpha, a + \alpha] \subset I$, et considérons $f([a - \alpha, a + \alpha]) = [b - \beta_1, b + \beta_2]$ avec $\beta_i > 0$ (le cas où a serait une extrémité de I se traiterait de manière analogue, à droite ou à gauche). Si k est un réel non nul de valeur absolue inférieure à $\min(\beta_1, \beta_2)$, étudions le rapport

$$\Delta(k) = \frac{g(b + k) - g(b)}{k}$$

Comme f et g sont des homéomorphismes réciproques, il existe un unique réel $h(k)$ dans $[-\alpha, \alpha] - \{0\}$ tel que $g(b + k) = a + h(k)$, soit $b + k = f(a + h(k))$. On a alors

$$\Delta(k) = \frac{h(k)}{f(a + h(k)) - f(a)}$$

Comme g est continue en b, on a $\lim\limits_{\substack{k \to 0 \\ k \neq 0}} h(k) = 0$, et le théorème sur la limite d'une composée de fonctions donne immédiatement

$$\lim_{\substack{k \to 0 \\ k \neq 0}} \Delta(k) = \frac{1}{f'(a)}$$

ce qui termine la démonstration. ■

EXEMPLE 8-4.13 Dérivées des fonction circulaires réciproques, des fonctions puissances :

- La fonction sin : $\left[-\frac{\pi}{2},\frac{\pi}{2}\right] \to [-1,1]$ est un homéomorphisme, dérivable en tout point, de dérivée non nulle sur $\left]-\frac{\pi}{2},\frac{\pi}{2}\right[$. Il en résulte que arcsin est dérivable en tout point de $]-1,1[$, avec pour tout $a \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[$

$$(\arcsin)'(\sin(a)) = \frac{1}{\cos a} = \frac{1}{\sqrt{1-\sin^2 a}}$$

puisque la fonction cos est positive sur $\left]-\frac{\pi}{2},\frac{\pi}{2}\right[$. On a donc finalement

$$\forall x \in]-1,1[\quad (\arcsin)'(x) = \frac{1}{\sqrt{1-x^2}}$$

- Comme $\arccos x = \frac{\pi}{2} - \arcsin x$, on obtient immédiatement

$$\forall x \in]-1,1[\quad (\arccos)'(x) = -\frac{1}{\sqrt{1-x^2}}$$

- De même la fonction tan : $\left]-\frac{\pi}{2},\frac{\pi}{2}\right[\to \mathbb{R}$ a une dérivée $\tan'(a) = 1+\tan^2 a$ non nulle en tout point. La fonction arctan est donc dérivable en tout point de $\mathbb{R}$, et comme $\forall a \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[$

$$(\arctan)'(\tan(a)) = \frac{1}{1+\tan^2 a}$$

on a

$$\forall x \in \mathbb{R} \quad (\arctan)'(x) = \frac{1}{1+x^2}$$

- Si $n \in \mathbb{N}^*$ la fonction $x \mapsto x^n$ définit un homéomorphisme $\mathbb{R}^{*+} \to \mathbb{R}^{*+}$, dont la dérivée ne s'annule pas. L'homéomorphisme réciproque $g : x \mapsto x^{\frac{1}{n}}$ est donc dérivable en tout point de $\mathbb{R}^{*+}$ et vérifie

$$\forall a > 0 \quad g'(a^n) = \frac{1}{na^{n-1}}$$

ce qui donne évidemment

$$\forall x > 0 \quad g'(x) = \frac{1}{n}x^{\frac{1}{n}-1}$$

- On en déduit aisément par composition que, pour $r \in \mathbb{Q}$ et f dérivable $I \to \mathbb{R}^{*+}$

$$(f^r)' = rf^{r-1}f'$$

EXERCICE 8-4.14 Retrouver à l'aide du théorème précédent les dérivées des fonctions hyperboliques réciproques (que l'on peut obtenir directement à partir des expressions faisant intervenir des logarithmes).

EXERCICE 8-4.15 Si f est une fonction numérique dérivable en a, avec $f(a) \neq 0$, on appelle **dérivée logarithmique de f** en a le rapport

$$\frac{f'(a)}{f(a)}$$

Montrer que, si f, g, h sont trois fonctions dérivables en a non nulles en a, la fonction

$$F = \frac{f^{\alpha}}{g^{\beta}h^{\gamma}}$$

possède une dérivée logarithmique en a égale à

$$\frac{F'(a)}{F(a)} = \alpha\frac{f'(a)}{f(a)} - \beta\frac{g'(a)}{g(a)} - \gamma\frac{h'(a)}{h(a)}$$

(On suppose pour le moment α, β, γ entiers si f, g, h sont à valeurs complexes, α, β, γ rationnels si f, g, h sont à valeurs réelles strictement positives). La formule précédente est parfois bien commode, notamment lorsque F est à valeurs réelles positives, car le signe de $F'(a)$ est alors le même que celui de l'expression précédente.

8-4.3 Dérivée d'ordre supérieur. Fonctions de classe C^p

8-4.3.1 Définition

Si $f : I \to E$ est dérivable au voisinage de a, on peut définir le fonction dérivée

$$f' : \left]a - \alpha, a + \alpha\right[\to E$$

pour $\alpha > 0$ suffisamment petit. Si cette fonction est-elle même dérivable en a, on appellera vecteur dérivé d'ordre 2 de f en a la limite

$$f''(a) = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f'(a+h) - f'(a)}{h}$$

et l'on pourra définir éventuellement des dérivées successives de f en a.

Supposer l'existence d'un vecteur dérivé d'ordre p pour f en a, que l'on notera $f^{(p)}(a)$, c'est déjà supposer implicitement l'existence des fonctions dérivées $f', f'', \ldots, f^{(p-1)}$ au voisinage de a.

DÉFINITION 8-4.16 *Si I est un intervalle de $\mathbb{R}$, $p \in \mathbb{N}^*$ et $f : I \to E$, on dit que f est de classe C^p sur I si f possède en tout point de I une dérivée d'ordre p et si la fonction*

$$f^{(p)} : I \to E$$

est continue sur I.

Comme une fonction continue sur I est dite de classe C^0 sur I, on a clairement

$$\forall p \in \mathbb{N}^* \ f \text{ est de classe } C^p \text{ sur } I \Leftrightarrow f \text{ est dérivable sur } I \text{ et } f' \text{ est de classe } C^{p-1}$$

Dans la définition précédente, si I contient une de ses extrémités, les dérivées successives en ce point sont des dérivées à droite ou à gauche.

On note $C^p(I,E)$ l'ensemble des fonctions de classe C^p de I vers E. Les propriétés de linéarité de la dérivation montrent aisément que $C^p(I,E)$ est un sous-espace vectoriel de $C^0(I,E)$. On notera de même

$$C^\infty(I,E) = \bigcap_{p\in\mathbb{N}} C^p(I,E)$$

l'espace des fonctions indéfiniment dérivables de I dans E, fonctions dites de classe C^∞.

8-4.3.2 Formule de Leibniz

THÉORÈME 8-4.17 Soient $f : I \to E_1$ et $g : I \to E_2$ admettant une dérivée d'ordre p en a. Si $\Phi : E_1 \times E_2 \to E$ est une application bilinéaire continue, la fonction $F : t \mapsto \Phi(f(t), g(t))$ admet une dérivée d'ordre p en a, et

$$F^{(p)}(a) = \sum_{k=0}^{p} \mathrm{C}_p^k \Phi\left(f^{(k)}(a), g^{(p-k)}(a)\right)$$

Démonstration : La démonstration se fait par récurrence sur p. Si $p=1$, il s'agit du théorème **8-4**.8 énoncé dans le cas d'une application bilinéaire. Supposons le résultat démontré pour les fonctions dérivables à l'ordre $p-1$. Comme f et g sont supposées posséder des dérivées d'ordre p en a, elles sont $p-1$ fois dérivables sur un voisinage de a de la forme $]a-\varepsilon, a+\varepsilon[$, avec $\varepsilon > 0$. Par hypothèse de récurrence, la fonction F est $p-1$ fois dérivable sur cet intervalle, et

$$\forall t \in]a-\varepsilon, a+\varepsilon[\quad F^{(p-1)}(t) = \sum_{k=0}^{p-1} \mathrm{C}_{p-1}^k \Phi\left(f^{(k)}(t), g^{(p-1-k)}(t)\right)$$

Le théorème **8-4**.8 montre alors que chacune des fonctions apparaissant dans la somme au second membre est dérivable en a, ce qui prouve que F est p fois dérivable en a. On a de plus

$$F^{(p)}(a) = \sum_{k=0}^{p-1} \mathrm{C}_{p-1}^k \left[\Phi\left(f^{(k+1)}(a), g^{(p-1-k)}(a)\right) + \Phi\left(f^{(k)}(a), g^{(p-k)}(a)\right)\right]$$

Par changement d'indice, on obtient facilement

$$F^{(p)}(a) = \Phi\left(f(a), g^{(p)}(a)\right) + \sum_{k=1}^{p-1} \left(\mathrm{C}_{p-1}^k + \mathrm{C}_{p-1}^{k-1}\right) \Phi\left(f^{(k)}(a), g^{(p-k)}(a)\right) + \Phi\left(f^{(p)}(a), g(a)\right)$$

ce qui donne le résultat à l'ordre p, compte tenu de la relation de récurrence permettant de calculer de proche en proche les coefficients du binôme. ■

Cette formule s'appliquera le plus souvent au produit de deux fonctions numériques, c'est-à-dire lorsque Φ est l'application $\mathbb{K}\times\mathbb{K} \to \mathbb{K}$ définie par $(x,y) \mapsto xy$. Elle est notamment utile pour dériver à un ordre élevé le produit d'une fonction polynôme par une fonction dont les dérivées successives sont bien connues :

EXEMPLE 8-4.18 Calculer la dérivée $n^{\text{ième}}$ de la fonction $x \mapsto (x^3+5x)\, e^{-2x}$.

COROLLAIRE 8-4.19 Si f et g sont de plus de classe C^p sur I, la fonction $\Phi(f,g)$ est de classe C^p. Cela est vrai en particulier pour le produit de deux fonctions numériques de $C^p(I,\mathbb{K})$ avec $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$. L'espace $C^p(I,\mathbb{K})$ est donc une sous-algèbre de $\mathbb{K}^I$.

COROLLAIRE 8-4.20 Si I et J sont deux intervalles de $\mathbb{R}$, $\varphi : I \to J$ et $f : J \to E$ sont de classe C^p, leur composée $f\circ\varphi$ est de classe C^p sur I.

Démonstration : Si $p=1$, la fonction $f\circ\varphi$ est dérivable sur I (théorème **8-4**.6), et sa fonction dérivée vaut $\varphi'.f'\circ\varphi$, fonction continue comme composée et produit de fonctions continues. Si le résultat est établi pour toutes fonctions de classe C^{p-1}, pour f et φ de classe C^p la fonction $f'\circ\varphi$ appartient à $C^{p-1}(I,E)$ comme composée de fonction C^{p-1}. La fonction φ' étant C^{p-1} sur I, le corollaire précédent (appliqué à $\Phi : \mathbb{R}\times E \to E$, $(\lambda,x)\mapsto \lambda x$) montre que le produit $(f\circ\varphi)' = \varphi'.f'\circ\varphi \in C^{p-1}(I,E)$, ce qui prouve finalement que la fonction $f\circ\varphi$ appartient à $C^p(I,E)$. ■

8-4.3.3 C^p-difféomorphisme d'intervalles

Nous anticiperons ici les résultats de la section suivante et admettrons qu'une fonction réelle dérivable sur un intervalle I, dont la dérivée est strictement positive sur cet intervalle, est strictement croissante sur I.

DÉFINITION 8-4.21 *Si I et J sont deux intervalles réels et $p\in\mathbb{N}^*$, une application $\varphi : I \to J$ est un C^p-difféomorphisme de I dans J ssi φ est bijective, de classe C^p et si la bijection réciproque $\varphi^{-1} : J : \to I$ est également de classe C^p.*

Le théorème suivant donne une caractérisation simple des C^p-difféomorphismes :

THÉORÈME 8-4.22 Si $\varphi : I \to \mathbb{R}$ est une fonction de classe C^p ($p\geqslant 1$), $J=\varphi(I)$ est un intervalle. Pour que φ induise un C^p-difféomorphisme de I dans J, il faut et il suffit que

$$\forall x\in I \qquad \varphi'(x)\neq 0$$

Démonstration : Si φ est C^p, elle est en particulier continue, et donc $\varphi(I) = J$ est un intervalle. Si $\varphi : I \to J$ est un C^p-difféomorphisme, c'est en particulier un homéomorphisme de I dans J. La fonction φ^{-1} est de plus dérivable en tout point de J, ce qui impose, d'après le théorème **8-4**.12, que la dérivée de φ ne s'annule pas sur I. Réciproquement, si φ est C^p sur I, avec φ' qui ne s'annule

pas sur I, le théorème des valeurs intermédiaires appliqué à la fonction continue φ' montre que cette fonction a un signe constant sur I. Nous reverrons dans la section suivante que cela entraîne que φ est strictement monotone sur I, et réalise donc, d'après le théorème **8-1**.9, un homéomorphisme $I \to J$. De plus, $\varphi^{-1} : J \to I$ est dérivable en tout point de J, avec

$$\forall y \in J \quad \left(\varphi^{-1}\right)'(y) = \frac{1}{\varphi'\left(\varphi^{-1}(y)\right)}$$

La fonction $\dfrac{1}{\varphi' \circ \varphi^{-1}}$ étant clairement continue sur J, φ^{-1} est au moins de classe C^1. Le théorème est donc bien démontré dans le cas où $p = 1$. Si on le suppose établi dans le cas C^{p-1} et si l'on suppose φ C^p, la fonction φ^{-1} est C^{p-1} par hypothèse de récurrence. Donc la fonction $\varphi' \circ \varphi^{-1}$ est C^{p-1} comme composée de telles fonctions, à valeurs dans $]0, +\infty[$ ou $]-\infty, 0[$. Il en est de même pour $\dfrac{1}{\varphi' \circ \varphi^{-1}}$, toujours à cause du théorème de composition, la fonction $x \to \dfrac{1}{x}$ étant C^∞ sur $]0, +\infty[$ ou $]-\infty, 0[$. La fonction φ^{-1} est donc bien C^p sur J, puisque sa dérivée est C^{p-1}. ■

8-4.3.4 Fonctions C^p par morceaux sur un segment

DÉFINITION 8-4.23 *Une fonction $f : [a,b] \to E$ est dite de classe C^p par morceaux sur ce segment si et seulement s'il existe une subdivision de $[a,b]$*

$$(d): \quad a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

telle que, sur chacun des intervalles ouverts de subdivision $]x_i, x_{i+1}[$, la restriction de f soit de classe C^p et soit prolongeable en une fonction de classe C^p sur $[x_i, x_{i+1}]$.

Cette définition est conforme à celle donnée pour $p = 0$ en **8-2**.2. On verra (théorème de prolongement d'une dérivée dans la section suivante) que cela signifie que f est de classe C^p sur tous les intervalles $]x_i, x_{i+1}[$, et que f ainsi que toutes ses dérivées jusqu'à l'ordre p possèdent des limites à droite et à gauche en tous les points de subdivision (à droite en a et à gauche en b). Lorsque l'espace $(E, \| \|)$ est complet (ce qui est le cas pour $E = \mathbb{R}$ ou $\mathbb{C}$), il suffira d'ailleurs que la dérivée d'ordre p de f possède des limites à droite et à gauche en ces points.

On vérifie aisément que l'espace $C^p_m([a,b], E)$ des fonctions C^p par morceaux sur $[a,b]$ est un espace vectoriel, qui contient l'espace $\mathcal{E}([a,b], E)$ des fonctions en escalier sur $[a,b]$.

8-5 Accroissements finis et applications

Nous étudierons d'abord le cas des fonctions réelles, et étendrons les résultats obtenus (avec quelques modifications) aux fonctions complexes ou vectorielles.

8-5.1 Cas des fonctions réelles

8-5.1.1 Théorème de Rolle

THÉORÈME 8-5.1 Si $f : [a,b] \to \mathbb{R}$ est une fonction continue, dérivable sur $]a,b[$ avec $f(a) = f(b)$, il existe un point $c \in]a,b[$ tel que $f'(c) = 0$.

Démonstration : Comme f est continue sur $[a,b]$, $f([a,b])$ est un segment $[m,M]$. Si $m = M$, la fonction f est constante sur $[a,b]$, donc sa dérivée est identiquement nulle. Si $m < M$, une de ces valeurs (au moins) est atteinte en un point c intérieur à $[a,b]$. Supposons par exemple $f(c) = m$, avec $c \in [a,b]$. Pour $x \in [a,b] - \{c\}$, on a

$$x < c \Rightarrow \frac{f(x) - f(c)}{x - c} \leqslant 0 \quad \text{et} \quad x > c \Rightarrow \frac{f(x) - f(c)}{x - c} \geqslant 0$$

En faisant tendre x vers c à droite et à gauche, on obtient $f'(c) \leqslant 0$ et $f'(c) \geqslant 0$, et donc $f'(c) = 0$. ■

On remarquera l'utilité de l'hypothèse $f(a) = f(b)$. Si elle n'est pas vérifiée, f peut atteindre ses extrema aux deux extrémités du segment.

EXERCICE 8-5.2 Si $P \in \mathbb{R}[X]$ est scindé, il en est de même de P' et de $P + \alpha P'$ pour $\alpha \in \mathbb{R}$ quelconque.

EXERCICE 8-5.3 Polynômes de Legendre sur $[a,b]$: on définit

$$L_n(X) = \frac{d^n}{dX^n}\left((X-a)^n (X-b)^n\right)$$

Montrer que L_n est un polynôme scindé, dont toutes les racines sont simples et sont dans $]a,b[$.

8-5.1.2 Formule des accroissements finis

THÉORÈME 8-5.4 Si $f : [a,b] \to \mathbb{R}$ est une fonction continue, dérivable sur $]a,b[$, il existe $c \in]a,b[$ tel que

$$f(b) - f(a) = (b-a) f'(c)$$

Démonstration : Il suffit d'appliquer le théorème de Rolle à la fonction

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a)$$

représentant la différence des ordonnées du point d'abscisse x sur la représentation graphique de f et du point de même abscisse sur la "corde" reliant les points $A(a, f(a))$ et $B(b, f(b))$. ■

L'interprétation géométrique de cette formule est claire si on l'écrit

$$\exists c \in]a,b[\qquad f'(c) = \frac{f(b) - f(a)}{b - a}$$

La tangente au point d'abscisse c à la représentation graphique de f est parallèle à la corde AB, puisque ces deux droites ont même coefficient directeur (figure **8.9**).

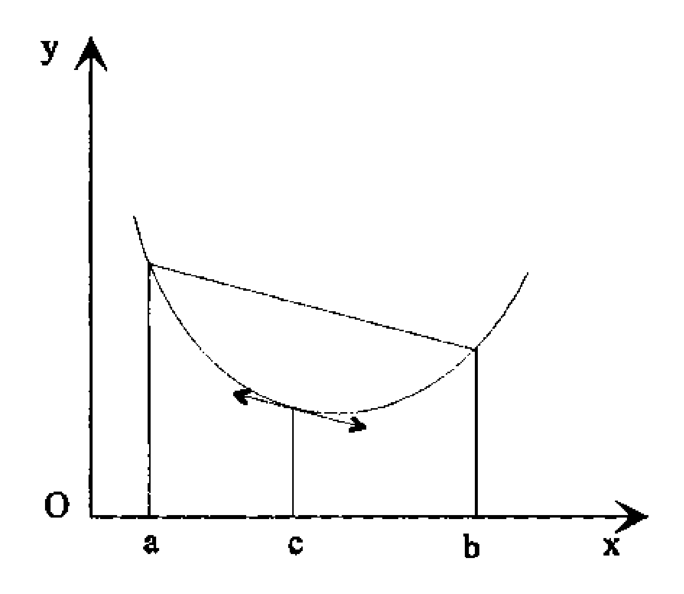


Figure **8.9** – *Théorème des accroissements finis*

L'existence du point c peut avoir son importance. Le théorème précédent est cependant le plus souvent utilisé sous la forme affaiblie suivante, qui restera valable pour les fonctions vectorielles :

COROLLAIRE 8-5.5 Si $f : [a,b] \to \mathbb{R}$ est une fonction continue, dérivable sur $]a,b[$, et s'il existe une constante M telle que, pour tout x de $]a,b[$, on ait $|f'(x)| \leqslant M$, on a alors

$$|f(b) - f(a)| \leqslant M(b-a)$$

8-5.1.3 Formule des accroissements finis généralisée

THÉORÈME 8-5.6 Si f et $g : [a,b] \to \mathbb{R}$ sont deux fonctions continues, dérivables sur $]a,b[$, il existe un $c \in]a,b[$ tel que

$$\begin{vmatrix} f(b) - f(a) & f'(c) \\ g(b) - g(a) & g'(c) \end{vmatrix} = 0$$

Démonstration : On remarque d'abord que, pour $g(x) = x$, on retrouve la formule des accroissements finis. Dans le cas général, il suffit d'appliquer le théorème de Rolle à la fonction

$$x \mapsto \begin{vmatrix} f(x) & f(a) & f(b) \\ g(x) & g(a) & g(b) \\ 1 & 1 & 1 \end{vmatrix}$$

(la nullité de ce déterminant traduit le fait que, dans le plan rapporté à un repère affine, les points $M(f(x), g(x))$, $A(f(a), g(a))$ et $B(f(b), g(b))$ sont alignés). On obtient l'existence d'un point c avec

$$\begin{vmatrix} f'(c) & f(a) & f(b) \\ g'(c) & g(a) & g(b) \\ 0 & 1 & 1 \end{vmatrix} = 0$$

ce qui donne le résultat souhaité. ∎

Ici encore, l'interprétation géométrique est simple. Si l'on suppose que

$$\forall x \in]a, b[\quad (f'(x), g'(x)) \neq (0, 0)$$

on peut considérer que $x \mapsto M(x)$, point du plan dont les coordonnées dans un repère $(O, \vec{\imath}, \vec{\jmath})$ sont $(f(x), g(x))$, définit un arc paramétré plan régulier : le vecteur dérivé

$$\vec{M}'(x) = f'(x)\,\vec{\imath} + g'(x)\,\vec{\jmath}$$

n'est jamais nul et dirige donc la tangente à l'arc au point de paramètre x. Le point $M(c)$ est alors tel que la tangente à l'arc en ce point est parallèle à la corde $M(a)M(b)$. Cette propriété des arcs plans ne se généralise pas en dimension supérieure, comme le montre l'exemple de l'hélice circulaire $t \mapsto (\cos t, \sin t, \alpha t)$ (repère orthonormé) sur $[0, 2\pi]$.

COROLLAIRE 8-5.7 Règle de l'Hospital : Si f et g sont des fonctions réelles continues, dérivables au voisinage de a sauf peut-être en a, la dérivée g' ne s'annulant pas sur ce voisinage (sauf éventuellement en a). Si le rapport des dérivées $\frac{f'}{g'}$ possède une limite en a (dans $\overline{\mathbb{R}}$), on a

$$\lim_{\substack{x \to a \\ \neq}} \frac{f(x) - f(a)}{g(x) - g(a)} = \lim_{\substack{x \to a \\ \neq}} \frac{f'(x)}{g'(x)}$$

Démonstration : Remarquons d'abord que l'hypothèse faite sur la dérivée g' montre que, pour $x \neq a$ voisin de a, la différence $g(x) - g(a)$ est non nulle (à cause de la formule des accroissements finis), ce qui permet de définir le rapport dont on étudie la limite. Supposons par exemple que la limite l du rapport $\frac{f'}{g'}$ soit finie (le cas $\pm\infty$ s'étudiant de façon analogue). Si $\epsilon > 0$ est donné, il existe $\alpha > 0$ tel que

$$\forall x \in]a - \alpha, a + \alpha[- \{a\} \qquad \left| \frac{f'(x)}{g'(x)} - l \right| < \epsilon$$

Pour $x \in]a - \alpha, a + \alpha[- \{a\}$, il existe c_x avec $0 < |c_x - a| < |x - a|$ tel que

$$\begin{vmatrix} f(x) - f(a) & f'(c_x) \\ g(x) - g(a) & g'(c_x) \end{vmatrix} = 0$$

d'après la formule des accroissements finis généralisée. Comme g' ne s'annule pas, cela peut s'écrire

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(c_x)}{g'(c_x)}$$

Comme $0<|c_x-a|<\alpha$, on a

$$0<|x-a|<\alpha \Rightarrow \left|\frac{f(x)-f(a)}{g(x)-g(a)}-l\right|<\varepsilon$$

ce qui donne bien

$$\lim_{\substack{x\to a\\ \neq}}\frac{f(x)-f(a)}{g(x)-g(a)}=l \quad \blacksquare$$

Deux remarques s'imposent ici :

- La règle de l'Hospital est d'un emploi extrêmement restreint, car très souvent on peut éviter d'y avoir recours, pour ne pas être obligé de rappeler ses hypothèses un peu pénibles. Par exemple, si f et g sont dérivables en a, il suffira très souvent d'écrire

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f(x)-f(a)}{x-a}\,\frac{x-a}{g(x)-g(a)}$$

pour voir que le rapport a pour limite $\dfrac{f'(a)}{g'(a)}$ lorsque ce rapport est défini. Plus généralement, on pourra trouver un équivalent de $f(x)-f(a)$ et de $g(x)-g(a)$ pour $x\to a$, par exemple à l'aide d'un développement limité, de la formule de Taylor-Young... pour lever l'indétermination.

- Il n'y a pas de réciproque : le rapport $\dfrac{f(x)-f(a)}{g(x)-g(a)}$ peut admettre une limite sans que $\dfrac{f'}{g'}$ en ait une. Prenons par exemple $g(x)=x$ et $f(x)=x^2\sin\dfrac{1}{x}$, prolongée par continuité en 0 par $f(0)=0$. On a

$$\lim_{\substack{x\to 0\\ \neq}}\frac{f(x)}{x}=0$$

et pourtant

$$\frac{f'(x)}{g'(x)}=f'(x)=2x\sin\frac{1}{x}-\cos\frac{1}{x}$$

n'a pas de limite en 0.

8-5.1.4 Application au sens de variation

THÉORÈME 8-5.8 Si $f : I \to \mathbb{R}$ est dérivable sur I, f est constante sur I si et seulement si f' est identiquement nulle sur cet intervalle.

Démonstration : La nullité de la dérivée est évidemment nécessaire pour que f soit constante. Réciproquement, si $f' \equiv 0$ sur I, pour $a, b \in I$, on a forcément $f(a) = f(b)$, à cause de la formule des accroissements finis. ■

THÉORÈME 8-5.9 Si $f : I \to \mathbb{R}$ est dérivable sur I, on a

- **f est croissante sur I si et seulement si $\forall x \in I \quad f'(x) \geqslant 0$.**
- **f est strictement croissante sur I si et seulement si sa dérivée est positive sur I et n'est identiquement nulle sur aucun sous-intervalle non réduit à un point de I.**

Démonstration : Si f est croissante sur I

$$\forall x, y \in I \quad x \neq y \Rightarrow \frac{f(x) - f(y)}{x - y} \geqslant 0$$

La définition de la dérivée comme limite du taux d'accroissement donne alors immédiatement $\forall x \in I \quad f'(x) \geqslant 0$. Réciproquement, si $f' \geqslant 0$ sur I, on a, pour $x < y$ dans I

$$\exists c \in]x, y[\quad f(y) - f(x) = (y - x) f'(c) \geqslant 0$$

ce qui montre que f est croissante.

Enfin, pour que f soit strictement croissante, il faut et il suffit évidemment qu'elle soit croissante et qu'elle ne soit constante sur aucun sous-intervalle de I non réduit à un point. ■

On remarquera que si l'implication $f' > 0$ sur $I \Rightarrow f$ strictement croissante est vraie, sa réciproque est évidemment fausse (penser par exemple à $x \mapsto x^3$ sur $\mathbb{R}$).

Dans le même ordre d'idée, la clé de la démonstration du théorème de Rolle était :

PROPOSITION 8-5.10 Si $f : I \to \mathbb{R}$ présente un extremum local en un point a intérieur à I et si f est dérivable en ce point, alors $f'(a) = 0$.

Ici encore l'exemple $x \mapsto x^3$ en $a = 0$ montre qu'il n'y a pas de réciproque. De même l'affirmation (pour une fonction f dérivable sur un intervalle ouvert) "f présente un extremum local en a donc f' s'annule en a en changeant de signe" est incorrecte. La fonction

$$f(x) = x^2 \cos^2 \frac{1}{x} + x^4 \sin^2 \frac{1}{x}$$

donne un contre-exemple en 0 : On a bien (exercice) $f'(0) = 0$, la fonction f présente un minimum absolu en 0. Pour $x \neq 0$, on a

$$f'(x) = 2x\cos^2\frac{1}{x} + 4x^3\sin^2\frac{1}{x} + (1-x^2)\sin\frac{2}{x}$$

Pour k entier on a facilement

$$f'\left(\frac{1}{\frac{\pi}{4}+k\frac{\pi}{2}}\right) \underset{k\to+\infty}{\sim} (-1)^k$$

ce qui montre que le signe de f' oscille au voisinage de 0.

8-5.2 Cas des fonctions vectorielles

Dans le cas de fonctions à valeurs dans un espace normé $(E, \| \ \|)$, la formule des accroissements finis n'est plus valable. On n'est plus assuré de l'existence d'un point c vérifiant

$$f(b) - f(a) = (b-a)f'(c)$$

Par exemple, pour $f : [0, 2\pi] \to \mathbb{C}$, définie par $f(t) = e^{it}$, on a $f(0) = f(2\pi)$, et pourtant la dérivée $f'(t) = ie^{it}$ ne s'annule jamais.

Cependant, la conséquence la plus importante de l'égalité des accroissements finis (corollaire **8-5**.5) est conservée dans le cas vectoriel : c'est **l'inégalité des accroissements finis.**

THÉORÈME 8-5.11 Si $f : [a,b] \to (E, \| \ \|)$ **est continue et dérivable sur** $]a,b[$**, et s'il existe une constante** $M \geqslant 0$ **telle que** $\forall t \in]a,b[\quad \|f'(t)\| \leqslant M$**, alors**

$$\|f(b) - f(a)\| \leqslant M(b-a)$$

Démonstration : On peut d'abord se ramener au cas où f est dérivable sur $[a,b]$. Si le théorème est démontré sous cette hypothèse, on obtiendra pour $0 < h < \frac{b-a}{2}$

$$\|f(b-h) - f(a+h)\| \leqslant M(b-a-2h)$$

ce qui donnera ensuite le résultat escompté en faisant tendre h vers 0, sous réserve effectivement de continuité de f en a et b. Comme de plus l'inégalité $\|f'(t)\| \leqslant M$ est large, elle permet de majorer le taux d'accroissement $\left\|\frac{f(t+h)-f(t)}{h}\right\|$ au voisinage de t par une constante plus grande que M. On commence donc par choisir $\varepsilon > 0$ et on remplace M par $M + \varepsilon$. On considère

$$X_\varepsilon = \{x \in [a,b] \mid \forall t \in [a,x] \quad \|f(t) - f(a)\| \leqslant (M+\varepsilon)(t-a)\}$$

X_ε est non vide puisque $a \in X_\varepsilon$. Il est clair que, de par sa définition, X_ε est un sous intervalle de $[a,b]$, qui contient sa borne supérieure, à

cause de la continuité de f. Pour obtenir l'inégalité escomptée, on veut montrer que cette borne supérieure vaut b. On fait un raisonnement par l'absurde, en supposant $X_\varepsilon = [a, c]$, avec $c < b$.

Comme $f'(c)$ existe, on peut trouver un réel $\alpha > 0$ tel que $\alpha < b - c$ et vérifiant

$$0 < h \leqslant \alpha \Rightarrow \left\| \frac{f(c+h) - f(c)}{h} - f'(c) \right\| \leqslant \varepsilon$$

ce qui entraîne, pour $0 < h \leqslant \alpha$

$$\|f(c+h) - f(c)\| \leqslant (\|f'(c)\| + \varepsilon) h \leqslant (M + \varepsilon) h$$

On arrive alors à une contradiction puisqu'on voit que $c + \alpha \in X_\varepsilon$. En effet, si $t \in [a, c + \alpha]$, on a soit $t \leqslant c$ donc

$$\|f(t) - f(a)\| \leqslant (M + \varepsilon)(t - a)$$

(puisque $c \in X_\varepsilon$), soit $t = c + h$ avec $0 < h \leqslant \alpha$ et on a alors également

$$\begin{aligned} \|f(t) - f(a)\| &\leqslant \|f(c+h) - f(c)\| + \|f(c) - f(a)\| \\ &\leqslant (M + \varepsilon)(h + c - a) = (M + \varepsilon)(t - a) \end{aligned}$$

On arrive donc finalement, puisque $b \in X_\varepsilon$

$$\|f(b) - f(a)\| \leqslant (M + \varepsilon)(b - a)$$

et donc $\|f(b) - f(a)\| \leqslant M(b - a)$ puisque l'inégalité précédente est valable pour tout $\varepsilon > 0$. ■

REMARQUE 8-5.12 Si l'on réfléchit à la démonstration précédente, on n'a véritablement utilisé que la dérivabilité à droite de f en c. Le résultat du théorème précédent est donc valable sous l'hypothèse de continuité de f sur $[a, b]$, et d'existence en tout point de $]a, b[$ d'une dérivée à droite de f majorée en norme par M. En raffinant un peu la démonstration précédente, on pourrait résoudre l'exercice suivant :

EXERCICE 8-5.13 Soient $f : [a, b] \to E$ et $g : [a, b] \to \mathbb{R}$ continues, dérivables en tout point de $]a, b[$ (si on veut, dérivables à droite) vérifiant

$$\forall t \in]a, b[\quad \|f'(t)\| \leqslant g'(t)$$

Montrer que

$$\|f(b) - f(a)\| \leqslant g(b) - g(a)$$

Interprétation cinématique ?

L'exercice précédent montre en particulier qu'une fonction numérique continue dérivable à droite en tout point d'un intervalle ouvert est croissante ssi la dérivée à droite est partout positive.

8-5.3 Applications des accroissements finis

8-5.3.1 Fonctions lipschitziennes dérivables

THÉORÈME 8-5.14 Si $f : I \to (E, \| \ \|)$ est une fonction dérivable et $K \in \mathbb{R}^+$, f est K-lipschitzienne sur I ssi

$$\forall t \in I \quad \|f'(t)\| \leqslant K$$

Démonstration : L'inégalité des accroissements finis montre que la condition est suffisante. Réciproquement, si f est K-lipschitzienne, pour $t, t' \in I$ avec $t' \neq t$, on a $\|f(t) - f(t')\| \leqslant K|t - t'|$ ce qui donne

$$\left\| \frac{f(t) - f(t')}{t - t'} \right\| \leqslant K$$

En faisant tendre t' vers t, par continuité de la norme on obtient $\|f'(t)\| \leqslant K$. ■

COROLLAIRE 8-5.15 Si $f : I \to E$ est dérivable, f est constante sur l'intervalle I si et seulement si f' est identiquement nulle sur I.

8-5.3.2 Théorème de prolongement d'une dérivée

THÉORÈME 8-5.16 Soient $a < b$ deux réels et $f :]a, b] \to E$ continue et dérivable sur $]a, b[$. Si la fonction f' possède une limite en b, alors f est dérivable (à gauche) en b et on a

$$f'(b) = \lim_{\substack{x \to b \\ <}} f'(x)$$

Démonstration : Si f est à valeurs réelles, on peut recopier la démonstration de la règle de l'Hospital pour montrer que

$$\lim_{x \to b} \frac{f(x) - f(b)}{x - b} = \lim_{x \to b} f'(x)$$

mais on utilise là l'égalité des accroissements finis, réservée aux fonctions réelles. Dans le cas général, on revient plutôt à la définition, et on montre, en appelant $l = \lim\limits_{x \to b} f'(x)$, que l'on a

$$\varphi(t) = f(t) - f(b) - (t - b)\,l \underset{t \to b}{=} o(t - b)$$

Si $\varepsilon > 0$ est fixé arbitrairement, on peut trouver $\alpha > 0$ tel que

$$\forall x \in [b - \alpha, b[\quad \|f'(x) - l\| < \varepsilon$$

Si t est choisi arbitrairement dans $[b - \alpha, b[$, la fonction φ est continue sur $[t, b]$, dérivable sur $]t, b[$ et vérifie sur cet intervalle

$$\|\varphi'(u)\| = \|f'(u) - l\| < \varepsilon$$

On a alors, par inégalité des accroissements finis

$$\forall t \in [b - \alpha, b[\quad \|\varphi(t)\| = \|\varphi(t) - \varphi(b)\| \leqslant \varepsilon(b - t)$$

ce qui prouve le résultat. ■

REMARQUE 8-5.17 Si $f :]a,b[\to E$ possède une limite l en b, il est assez logique de "poser" $f(b) = l$. On prolonge ainsi par continuité la fonction f en b, et c'est ce prolongement qui est naturel. Lorsque $f :]a,b] \to E$ est continue, dérivable sur $]a,b[$, la dérivée f' ayant une limite l' en b, dire

$$\text{" on pose } f'(b) = l' \text{ "}$$

est incorrect, car l'utilisation du symbole $f'(b)$ suppose la dérivabilité de f en b. C'est le théorème précédent qu'il faut invoquer, et il est beaucoup plus satisfaisant de dire : " le théorème de prolongement d'une dérivée assure que f est dérivable en b avec $f'(b) = l'$ ".

REMARQUE 8-5.18 Si $(E, \| \|)$ est complet (ce qui est le cas usuellement avec $E = \mathbb{R}$ ou $\mathbb{C}$), si on suppose simplement f continue et dérivable sur $]a,b[$, la dérivée f' ayant une limite l en b, alors f peut-être prolongée par continuité en b, en une fonction qui sera dérivable en b d'après le théorème précédent. En effet, f' ayant une limite en b est bornée au voisinage de b :

$$\exists M > 0 \quad \exists \alpha > 0 \quad \forall t \in]b-\alpha, b[\quad \|f'(t)\| \leqslant M$$

La fonction f est alors M-lipschitzienne sur $]b-\alpha, b[$, donc est uniformément continue sur cet intervalle. Elle possède alors, si $(E, \| \|)$ est complet, une limite L en b (théorème de prolongement d'une application uniformément continue). En prolongeant f par continuité en b par $f(b) = L$, on est alors ramené aux hypothèses du théorème de prolongement d'une dérivée.

REMARQUE 8-5.19 En appliquant le théorème de prolongement d'une dérivée à $f^{(p-1)}$, puis à $f^{(p-2)}$ etc... jusqu'à f, sur chacun des intervalles de subdivision, on obtient aisément la caractérisation des fonctions de classe C^p par morceaux sur un segment $[a,b]$:

$f : [a,b] \to E$ est dans $C^p_m([a,b], E)$ ssi on peut trouver une subdivision (x_i) de $[a,b]$ telle que

- $\forall i \quad f$ est de classe C^p sur $]x_i, x_{i+1}[$
- f et toutes ses dérivées d'ordre $\leqslant p$ possèdent des limites à droite et à gauche aux points de subdivision.

Si l'espace est complet, la remarque précédente montre qu'il suffit de supposer l'existence de limites pour la dérivée d'ordre p.

EXERCICE 8-5.20 Montre que la fonction définie sur $\mathbb{R}$ par

$$\begin{cases} f(t) = e^{-\frac{1}{t^2}} & \text{si} \quad t > 0 \\ f(t) = 0 & \text{si} \quad t \leqslant 0 \end{cases}$$

est de classe C^∞ sur $\mathbb{R}$. Indication : f est évidemment de classe C^∞ sur $\mathbb{R}^{*+}$ et $\mathbb{R}^{*-}$, et on peut montrer que

$$\forall p \in \mathbb{N}^* \quad \forall t > 0 \quad f^{(p)}(t) = e^{-\frac{1}{t^2}} P_p\left(\frac{1}{t}\right)$$

où P_p est une fonction polynôme.

8-5.3.3 Inégalité de Taylor Lagrange

THÉORÈME 8-5.21 Soit $f : I \to (E, \| \|)$ **une fonction de classe** C^{p+1}**. Si** a **et** b **sont deux points de** I **et** $K = [a,b]$ **ou** $[b,a]$ **le segment d'extrémités** a **et** b**, on a**

$$\left\| f(b) - f(a) - \sum_{k=1}^{p} \frac{(b-a)^k}{k!} f^{(k)}(a) \right\| \leqslant \frac{|b-a|^{p+1}}{(p+1)!} \sup_{x \in K} \left\| f^{(p+1)}(x) \right\|$$

Démonstration : Remarquons que les hypothèses entraînant la continuité de $f^{(p+1)}$ sur le segment K, cette fonction est bornée sur K. Lorsque E est de dimension finie (ou est complet), la démonstration la plus rapide consiste à utiliser la formule de Taylor avec reste sous forme d'intégrale, puisque f est C^{p+1} sur le segment K

$$\Delta_p = f(b) - f(a) - \sum_{k=1}^{p} \frac{(b-a)^k}{k!} f^{(k)}(a) = \int_a^b \frac{(b-t)^p}{p!} f^{(p+1)}(t)\, dt$$

l'inégalité de la norme donnant alors (voir chapitre d'intégration)

$$\|\Delta_p\| \leqslant \left| \int_a^b \frac{|b-t|^p}{p!} \left\| f^{(p+1)}(t) \right\| dt \right| \leqslant \left| \int_a^b \frac{|b-t|^p}{p!} \left\| f^{(p+1)} \right\|_{\infty,K} dt \right|$$
$$= \frac{|b-a|^{p+1}}{(p+1)!} \sup_{x \in K} \left\| f^{(p+1)}(x) \right\|$$

Sous des hypothèses moins fortes, on peut préférer la démonstration suivante :

La fonction φ définie par

$$\varphi(t) = f(b) - f(t) - \sum_{k=1}^{p} \frac{(b-t)^k}{k!} f^{(k)}(t)$$

est par hypothèse de classe C^1 sur K (somme et produits de fonctions de classe C^1 sur K) avec par un calcul simple.

$$\varphi'(t) = -\frac{(b-t)^p}{p!} f^{(p+1)}(t)$$

et donc, en posant $M = \sup_{x \in K} \left\| f^{(p+1)}(x) \right\|$

$$\|\varphi'(t)\| \leqslant \frac{|b-t|^p}{p!} M = g'(t)$$

où g est la fonction définie sur K par $g(t) = \pm M \dfrac{(b-t)^{p+1}}{(p+1)!}$ où le signe $\pm$ est fixé sur K en fonction des positions respectives de a et b et de la parité de p. Le résultat de l'exercice **8-5.**13 donne alors

$$\|\varphi(b) - \varphi(a)\| \leqslant |g(b) - g(a)|$$

ce qui est l'inégalité souhaitée. ■

REMARQUE 8-5.22 La seconde démonstration qui précède montre que le résultat du théorème est valable lorsque f est supposée de classe C^p sur $[a,b]$ et $p+1$ fois dérivable sur $]a,b[$, avec une dérivée d'ordre $p+1$ bornée sur $]a,b[$.

Lorsque la fonction f est **à valeurs réelles**, on peut obtenir le raffinement suivant (analogue à l'**égalité** des accroissements finis). Le résultat est intéressant, mais de portée assez limitée, puisqu'on l'affaiblit souvent sous forme d'une inégalité.

THÉORÈME 8-5.23 Si $f : [a,b]$ (ou $[b,a]$) $\to \mathbb{R}$ est de classe C^p et possède une dérivée d'ordre $p+1$ sur $]a,b[$, il existe un point $c \in]a,b[$ tel que

$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f''(a) + \cdots + \frac{(b-a)^p}{p!}f^{(p)}(a) + \frac{(b-a)^{p+1}}{(p+1)!}f^{(p+1)}(c)$$

Démonstration : Il suffit d'appliquer le théorème des accroissements finis sur $[a,b]$ à la fonction φ définie dans la démonstration du théorème précédent. ■

EXERCICE 8-5.24 Calculer $\sin(31°)$ avec une erreur inférieure à 10^{-5}.

Il faut penser à utiliser l'égalité des accroissements finis ou de Taylor pour démontrer des inégalités globales entre fonctions, ce qui permet d'éviter le recours systématique à l'étude des variations d'une fonction auxiliaire (méthode efficace mais qui fait souvent perdre du temps).

EXERCICE 8-5.25 Montrer que, pour tout $x > 0$

$$x - \frac{x^2}{2} < \ln(1+x) < x$$
$$\frac{x}{1+x^2} < \arctan x < x$$

8-6 Développements limités

On travaille dans cette section avec des fonctions définies au voisinage d'un point $x_0 \in \mathbb{R}$ privé éventuellement de x_0 (avec éventuellement $x_0 = +\infty$) à valeurs dans un espace normé $(E, \|\ \|)$ qui est le plus souvent le corps de base $\mathbb{R}$ ou $\mathbb{C}$. Ce voisinage est parfois voisinage à gauche (ou à droite).

8-6.1 Généralités

DÉFINITION 8-6.1 *$f : \mathcal{V}(0) \ni V \to E$ étant donnée, et $n \in \mathbb{N}$ étant un entier naturel, on dit que f possède un développement limité d'ordre n au voisinage de 0 ssi il existe des vecteurs $(a_0, a_1, \ldots, a_n) \in E^{n+1}$ tels que*

$$f(x) \underset{0}{=} a_0 + xa_1 + \cdots + x^n a_n + o(x^n)$$

Si f est définie au voisinage de x_0, on dira que f admet un développement limité d'ordre n en x_0 ssi la fonction $h \mapsto f(x_0 + h)$ admet un développement du même ordre en 0. On aura alors une égalité de la forme

$$f(x) \underset{x_0}{=} a_0 + (x - x_0)a_1 + \cdots + (x - x_0)^n a_n + o((x - x_0)^n)$$

Enfin, si f est définie au voisinage de $+\infty$ (ou $-\infty$), on se ramènera au voisinage de 0 en travaillant avec $x \mapsto f\left(\frac{1}{x}\right)$. On aura alors

$$f(x) \underset{+\infty}{=} a_0 + \frac{1}{x}a_1 + \cdots + \frac{1}{x^n}a_n + o\left(\frac{1}{x^n}\right)$$

Au voisinage de x_0 on travaille avec "l'infiniment petit principal" $(x - x_0)$, au voisinage de $\pm\infty$ c'est $\dfrac{1}{x}$ qui joue ce rôle. Dans tous les cas, pour "calculer" des DL, on se ramène par changement de variable au voisinage de 0. On notera en abrégé $DL^n_{x_0}(f)$ pour dire "développement limité d'ordre n de f en x_0", toute utilisation de ce symbole supposant donc f définie au voisinage de x_0.

Attention ! *Nous donnons ici la définition* ***d'un développement limité polynomial****. Il est bien évident que, pour étudier par exemple au voisinage de 0 la fonction définie sur $\mathbb{R}^{*+}$ par $f(x) = x^x$, on écrira*

$$f(x) = e^{x \ln x}$$

et que, dans ce cas, c'est $u = x \ln x$ qui jouera le rôle d'infiniment petit "principal" (on a bien $\lim_0 u(x) = 0$), et que le développement

$$f(x) = 1 + x \ln x + \frac{(x \ln x)^2}{2} + o(x \ln x)^2$$

(obtenu à partir du développement polynomial de la fonction exponentielle en 0) jouera un rôle identique à celui d'un développement limité polynomial. On parle alors plutôt de développement asymptotique.

THÉORÈME 8-6.2 Si f admet un DL^n_0, le "polynôme" (à coefficients dans E)

$$P_n(x) = a_0 + x a_1 + \cdots + x^n a_n$$

est unique, on l'appelle la partie régulière du $DL^n_0(f)$. Si ce polynôme est non nul et si p est sa valuation ($a_p x^p$ est le terme de plus bas degré de P_n, tel que $a_p \neq 0$) on a alors

$$f(x) \underset{0}{\sim} x^p a_p$$

Démonstration : Si on a deux polynômes de degrés inférieurs ou égaux à n tels que

$$f(x) = P_n(x) + o(x^n) = Q_n(x) + o(x^n)$$

on a alors, puisque la classe $o(x^n)$ est stable par combinaisons linéaires,

$$P_n(x) - Q_n(x) = o(x^n)$$

Si $P_n \neq Q_n$, le polynôme $P_n - Q_n$ non nul est de valuation p avec $0 \leqslant p \leqslant n$ et est équivalent en 0 à son terme de plus bas degré $x^p\, b_p$ avec $b_p \neq 0_E$. On doit avoir

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{x^p\, b_p}{x^n} = 0_E$$

ce qui n'est évidemment pas possible. ■

ATTENTION ! Lorsqu'on utilise un DL, il est indispensable d'utiliser le symbole = et en conséquence d'écrire le "terme complémentaire" $o(x^n)$, qu'on peut également écrire $x^n\varepsilon(x)$ avec $\lim\limits_{x\to 0} \varepsilon(x) = 0_E$. Il n'est pas question de "négliger" ce qui est "négligeable". Par exemple écrire

$$\tan x \underset{0}{=} x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5)$$

peut se traduire, si l'on tient à tout prix à utiliser le symbole "$\sim$"

$$\begin{cases} \tan x \sim x & \text{soit} \quad \lim\limits_{x\to 0} \dfrac{\tan x}{x} = 1 \\ \tan x - x \sim \dfrac{x^3}{3} & \text{soit} \quad \lim\limits_{x\to 0} \dfrac{\tan x - x}{x^3} = \dfrac{1}{3} \\ \tan x - x - \dfrac{x^3}{3} \sim \dfrac{2x^5}{15} & \text{soit} \quad \lim\limits_{x\to 0} \dfrac{\tan x - x - \frac{x^3}{3}}{x^5} = \dfrac{2}{15} \end{cases}$$

chaque ligne précisant la ligne précédente. Ecrire

$$\tan x \underset{0}{\sim} x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5) \text{ ou } \tan x \underset{0}{\sim} x + \frac{x^3}{3} + \frac{2x^5}{15}$$

fait perdre l'information contenue dans le développement d'ordre 5 et se résume à $\tan x \sim x$. Ecrire

$$\tan x \underset{0}{=} x + \frac{x^3}{3} + \frac{2x^5}{15}$$

est incorrect puisque la fonction $x \mapsto \tan x$ n'est pas polynomiale ! Ces dernières écritures ont de plus le redoutable avantage d'amener beaucoup d'erreurs et de résultats faux notamment lorsqu'on utilise les DL pour lever des indéterminations. Elles sont donc à proscrire.

COROLLAIRE 8-6.3 (Troncature) Si f admet un DL_0^n, elle admet pour tout entier $p \leqslant n$ un DL_0^p dont la partie régulière est obtenue en ne gardant dans la partie régulière du $DL_0^n(f)$ que les termes de degrés inférieurs ou égaux à p. On dit que le $DL_0^p(f)$ est déduit par troncature à l'ordre p du $DL_0^n(f)$.

COROLLAIRE 8-6.4 (Parité et imparité) Si f est une fonction paire définie au voisinage de 0 et possède un DL_0^n, la partie régulière de ce développement limité est un polynôme pair (donc ne contient que des monômes de degrés pairs). On a un résultat analogue pour les fonctions impaires.

Démonstration : Si f est paire et vérifie $f(x) = P_n(x) + o(x^n)$, on a , en composant à droite ("changement de variable") avec la fonction $x \mapsto -x$ **qui tend vers 0** en 0,

$$f(x) = f(-x) = P_n(-x) + o(x^n)$$

puisque les classes $o(x^n)$ et $o((-x)^n)$ sont évidemment égales. Par unicité de la partie régulière du DL_0^n, les polynômes $P_n(X)$ et $P_n(-X)$ sont donc (formellement) égaux, ce qui donne le résultat. ■

Le théorème d'unicité n'a pas de réciproque : deux fonctions définies au voisinage de zéro peuvent avoir des développements limités à tous ordres avec mêmes parties régulières sans être égales. Par exemple la fonction

$$f(x) = e^{-\frac{1}{x^2}} \text{ avec } f(0) = 0$$

admet un DL_0^n pour tout n qui s'écrit

$$f(x) = o(x^n)$$

par croissance comparée au voisinage de $+\infty$ de e^{-u} et de $\dfrac{1}{u^k}$ (avec le "changement de variable" $u = \dfrac{1}{x^2}$). Ce développement limité ne permet pas de distinguer f de la fonction identiquement nulle.

8-6.2 Développement limité et dérivabilité

THÉORÈME 8-6.5 Si f admet un DL_0^n avec $n \geqslant 1$,

$$f(x) = a_0 + x\,a_1 + \cdots + x^n\,a_n + o(x^n)$$

alors la fonction f (éventuellement prolongée par continuité en 0 en posant $f(0) = a_0$) est dérivable en 0 et vérifie $f'(0) = a_0$. Dans le cas $E = \mathbb{R}$, la droite d'équation

$$y = a_0 + a_1 x$$

est donc la tangente à la représentation graphique de la fonction f au point d'abscisse 0.

On a un résultat analogue pour une étude au voisinage d'un point x_0 quelconque.

Ce théorème est le seul résultat de régularité pour la fonction f qu'on puisse déduire *a priori* de l'existence d'un développement en 0. En particulier, on n'a aucun renseignement en ce qui concerne la continuité de f sur un voisinage de 0, donc *a fortiori* sur la dérivabilité de f en dehors de 0, et donc sur l'existence d'une dérivée seconde de f en 0 si $n \geqslant 2$. La fonction

$$f(x) = x + x^4 \mathbf{1}_{\mathbb{Q}}(x)$$

où $\mathbf{1}_{\mathbb{Q}}$ est la fonction caractéristique des rationnels est discontinue partout sauf à l'origine (tout réel est limite d'une suite de rationnels et d'une suite d'irrationnels), mais présente un DL_0^3 qui s'écrit

$$f(x) = x + o(x^3)$$

8-6.3 "Intégration" d'un développement limité

THÉORÈME 8-6.6 On suppose f dérivable au voisinage de 0, avec f' admettant un DL_0^n, donné par

$$f'(x) = a_0 + x\,a_1 x + \cdots + x^n\,a_n + o(x^n)$$

alors f admet un DL_0^{n+1} donné par

$$f(x) = f(0) + x\,a_0 + x^2\frac{a_1}{2} + \cdots + x^{n+1}\frac{a_n}{n+1} + o(x^{n+1})$$

dont la partie régulière est donnée en intégrant la partie régulière du développement de f' (sans oublier la "constante d'intégration" $f(0)$!).

Démonstration : La fonction

$$\varphi(t) = f(t) - \left(f(0) + ta_0 + t^2\frac{a_1}{2} + \cdots + t^{n+1}\frac{a_n}{n+1}\right)$$

est dérivable au voisinage de 0 et vérifie par hypothèse $\varphi'(t) \underset{0}{=} o(t^n)$. Si $\varepsilon > 0$ est fixé, il existe donc $\alpha > 0$ tel que

$$\forall\, t \in [-\alpha, \alpha] \quad \|\varphi'(t)\| \leqslant \varepsilon\, |t|^n$$

Si $x \in [-\alpha, \alpha]$, on a donc sur $[0, x]$ (ou $[x, 0]$) $\|\varphi'(t)\| \leqslant \varepsilon\,|x|^n$. L'inégalité des accroissements finis montre alors que

$$|x| \leqslant \alpha \Rightarrow \|\varphi(x) - \varphi(0)\| = \|\varphi(x)\| \leqslant \varepsilon\,|x|^{n+1}$$

ce qui démontre bien que

$$f(x) = f(0) + x.a_0 + x^2\frac{a_1}{2} + \cdots + x^{n+1}\frac{a_n}{n+1} + o(x^{n+1}) \quad \blacksquare$$

Attention, il n'y a pas de réciproque à ce théorème, l'existence d'un DL_0^n pour une fonction dérivable f n'entraîne pas forcément l'existence d'un DL_0^{n-1} pour la dérivée f'. Par exemple, la fonction

$$f(x) = x^3 \sin\left(\frac{1}{x^2}\right) \text{ prolongée par } f(0) = 0$$

est dérivable en tout point de $\mathbb{R}^*$ comme composée de fonctions dérivables. Elle est aussi dérivable à l'origine et $f'(0) = 0$ puisque $f(x) \underset{0}{=} o(x^2)$. On a donc un $DL_0^2(f)$. Pour $x \neq 0$ on a

$$f'(x) = 3x^2 \sin\left(\frac{1}{x^2}\right) - 2\cos\left(\frac{1}{x^2}\right)$$

et il est facile de voir que cette fonction n'ayant pas de limite en 0 ne peut avoir un DL_0^1.

REMARQUE 8-6.7 Cependant, si l'on sait que f' possède un développement limité d'ordre $n-1$, la partie régulière peut être obtenue en dérivant la partie régulière du développement limité d'ordre n de f puisqu'on sait que cette dernière est primitive du polynôme cherché.

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \frac{x^{n+1}}{1-x}$$

donne immédiatement :

$$\frac{1}{1-x} \underset{0}{=} 1 + x + x^2 + \cdots + x^n + o(x^n)$$

En changeant x en $-x$ (inf.petit avec x) on a :

$$\frac{1}{1+x} \underset{0}{=} 1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)$$

Le changement de x en x^2 (inf.petit avec x) donne :

$$\frac{1}{1+x^2} \underset{0}{=} 1 - x^2 + x^4 + \cdots + (-1)^n x^{2n} + o(x^{2n}).$$

En intégrant ces développements limités :

$$\ln(1-x) \underset{0}{=} -x - \frac{x^2}{2} - \cdots - \frac{x^{n+1}}{n+1} + o(x^{n+1})$$

$$\ln(1+x) \underset{0}{=} x - \frac{x^2}{2} - \cdots + (-1)^n \frac{x^{n+1}}{n+1} + o(x^{n+1})$$

$$\arctan x \underset{0}{=} x - \frac{x^3}{3} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+1})$$

On obtiendrait de même

$$\operatorname{argth}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) = x + \frac{x^3}{3} + \cdots + \frac{x^{2n+1}}{2n+1} + o(x^{2n+1})$$

soit en intégrant le DL_0^{2n} de la dérivée $\frac{1}{1-x^2}$ soit en combinant les DL de $\ln(1-x)$ et $\ln(1+x)$.

EXERCICE 8-6.8 En admettant que $x \mapsto \frac{1}{(1-x)^{p+1}}$ avec $p \in \mathbb{N}$ possède un DL_0^n démontrer que

$$\frac{1}{(1-x)^{p+1}} \underset{0}{=} 1 + \mathrm{C}_{p+1}^1 x + \cdots + \mathrm{C}_{p+n}^n x^n + o(x^n)$$

(l'existence de ce DL_0^n peut être justifiée par le théorème de Taylor-Young, ou par le théorème concernant le produit de développements limités).

8-6.4 Théorème de Taylor-Young

Ce résultat est très souvent utilisé pour justifier l'existence de développements limités. Hormis le cas des fonctions "usuelles", cette "formule" est rarement utilisée pour calculer explicitement des développements limités. On l'utilise plutôt à l'inverse pour, connaissant un développement limité de f en x_0, obtenir sans calculs supplémentaires les valeurs des dérivées successives de f en x_0.

THÉORÈME 8-6.9 (Taylor-Young) Si f est définie au voisinage de x_0 et possède une dérivée d'ordre n en x_0 (ce qui suppose l'existence de dérivées d'ordre inférieur au voisinage de x_0), elle admet un développement limité d'ordre n en x_0 donné par

$$f(x) \underset{x_0}{=} f(x_0) + (x - x_0)f'(x_0) + \cdots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0) + o\left((x - x_0)^n\right)$$

Démonstration : Par translation on peut se ramener en 0. Pour $n = 1$, l'écriture

$$f(x) = f(0) + xf'(0) + o(x)$$

correspond exactement à la définition de la dérivée $f'(0)$. On procède ensuite par récurrence sur n. On suppose que toute fonction définie au voisinage de 0 vérifiant les hypothèses du théorème à l'ordre $n - 1$ admet un développement d'ordre $n-1$ donné par la formule annoncée. Si f vérifie les hypothèses à l'ordre n, la fonction f' les vérifie à l'ordre $n - 1$ et on a donc

$$f'(x) = f'(0) + xf''(0) + \cdots + \frac{x^{n-1}}{(n-1)!} f^{(n)}(0) + o(x^{n-1})$$

Le théorème d'intégration permet alors de conclure immédiatement. ■

Attention ! *Il ne faut pas se méprendre sur la ressemblance de cette "formule" de Taylor Young avec le résultat du théorème de Taylor-Lagrange (section* **8-5.3.3***). La formule de Taylor-Young donne le comportement* ***asymptotique*** *de la différence*

$$\Delta(x) = f(x) - \left(f(x_0) + (x - x_0)f'(x_0) + \cdots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)\right)$$

pour x tendant vers x_0. On peut bien sûr en déduire une majoration du type

$$\|\Delta(x)\| \leqslant \varepsilon\, |x - x_0|^n$$

au voisinage de x_0, mais ***sans information précise sur la taille de ce voisinage****. L'inégalité de Taylor-Lagrange donne aussi une majoration de $\|\Delta(x)\|$, mais pour tout x d'un segment $[x_0, x_1]$ par exemple, à partir* ***d'hypothèses globales*** *sur la dérivée $n + 1^{\text{ème}}$ de la fonction f sur cet intervalle. Un simple développement limité ne donne pas une telle information.*

EXERCICE 8-6.10 Si $P_n(x)$ est la partie régulière du développement limité de la fonction tan à l'ordre n en 0, montrer que

$$\forall x \in \left]0, \frac{\pi}{2}\right[\quad \tan x > P_n(x)$$

En supposant connues les propriétés de dérivabilité des fonctions usuelles, la formule de Taylor-Young donne en particulier :

$$\sin x \underset{0}{=} x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2})$$

$$\cos x \underset{0}{=} 1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1})$$

$$e^x \underset{0}{=} 1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} + o(x^n)$$

$$\operatorname{sh} x \underset{0}{=} x + \frac{x^3}{6} + \frac{x^5}{120} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2})$$

$$\operatorname{ch} x \underset{0}{=} 1 + \frac{x^2}{2} + \frac{x^4}{24} + \cdots + \frac{x^{2n}}{(2n)!} + o(x^{2n+1})$$

$$(1+x)^\alpha \underset{0}{=} 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + o(x^n)$$

où α est un réel fixé.

EXERCICE 8-6.11 Quels sont les DL_0^{2n} des fonctions $x \mapsto \dfrac{1}{\sqrt{1-x^2}}$ et $x \mapsto \dfrac{1}{\sqrt{1+x^2}}$? En déduire que

$$\arcsin x = \sum_{k=0}^{n} \frac{1}{2^{2k}(2k+1)} \mathrm{C}_{2k}^{k} x^{2k+1} + o(x^{2n+2})$$

$$\operatorname{argsh} x = \ln\left(x + \sqrt{1+x^2}\right) = \sum_{k=0}^{n} \frac{(-1)^k}{2^{2k}(2k+1)} \mathrm{C}_{2k}^{k} x^{2k+1} + o(x^{2n+2})$$

8-6.5 Opérations et développements limités

8-6.5.1 Somme et combinaison linéaire

THÉORÈME 8-6.12 Si f et $g : \mathcal{V}(0) \ni V \to E$ admettent toutes deux un DL_0^n et si α et β sont deux scalaires, alors $\alpha f + \beta g$ admet un DL_0^n dont la partie régulière est la combinaison linéaire correspondante des parties régulières :

$$f(x) = P_n(x) + o(x^n) \text{ et } g(x) = Q_n(x) + o(x^n) \Longrightarrow$$
$$(\alpha f + \beta g)(x) = \alpha P_n(x) + \beta Q_n(x) + o(x^n)$$

Il est très important de noter que l'on travaille avec des développements limités **de même ordre** pour f et g. Si on dispose d'un $DL_0^n(f)$ et d'un $DL_0^p(g)$ avec $p < n$, les monômes de degrés strictement supérieurs à p de la partie régulière du $DL_0^n(f)$ seront non significatifs et devront être traités comme $o(x^p)$ lorsqu'on fera un développement limité de $\alpha f + \beta g$. Dans le cas général, on ne pourra obtenir mieux qu'un $DL_0^p(\alpha f + \beta g)$.

Les développements limités des fonctions $x \mapsto \operatorname{sh} x$ et $x \mapsto \operatorname{ch} x$ peuvent être obtenus ainsi à partir du développement de la fonction $x \mapsto e^x$.

8-6.5.2 Produit

On suppose ici les fonctions numériques à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. On pourrait généraliser le résultat à l'aide de la notion d'application bilinéaire continue entre espaces normés.

THÉORÈME 8-6.13 Si f et $g : \mathcal{V}(0) \ni V \to \mathbb{K}$ admettent toutes deux un DL_0^n donnés par

$$f(x) = P_n(x) + o(x^n) \text{ et } g(x) = Q_n(x) + o(x^n)$$

alors le produit $x \mapsto f(x)g(x)$ admet un DL_0^n dont la partie régulière est obtenue en tronquant le polynôme P_nQ_n à l'ordre n.

Démonstration : Les fonctions polynômes P_n et Q_n étant bornées au voisinage de 0, leurs produits avec des fonctions de la classe $o(x^n)$ sont encore dans cette classe. On a donc

$$f(x)g(x) = P_n(x)Q_n(x) + o(x^n)$$

et on ne retient dans P_nQ_n que les monômes de degré $\leqslant n$, les autres étant dans $o(x^n)$. ∎

EXERCICE 8-6.14 Donner $DL_0^3\left(\dfrac{e^x}{1-x}\right)$ (réponse : $1 + 2x + \frac{5}{2}x^2 + \frac{8}{3}x^3 + o\left(x^3\right)$).

REMARQUE 8-6.15 Si f est infiniment petit d'ordre p et g est infiniment petit d'ordre q, la connaissance d'un $DL_0^{n+p}(f)$ et d'un $DL_0^{n+q}(g)$ permet d'obtenir un développement limité d'ordre $n+p+q$ pour fg, puisqu'on peut, d'après le théorème précédent, obtenir un développement d'ordre n du produit $\dfrac{f(x)}{x^p} \cdot \dfrac{g(x)}{x^q}$. Par exemple, la connaissance d'un développement d'ordre 5 pour $x \mapsto \sin x$ permet l'obtention d'un développement d'ordre 7 pour $x \mapsto \sin^3 x$ (7=4+3).

8-6.5.3 Composition

On compose ici une fonction à valeurs dans $\mathbb{R}$ avec une fonction vectorielle. Il est essentiel de bien noter les hypothèses du résultat suivant, qu'on peut résumer de manière imprécise par la nécessité de travailler avec des "infiniment petits" :

THÉORÈME 8-6.16 Soit f définie au voisinage de 0, à valeurs dans $\mathbb{R}$, admettant un DL_0^n :

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n + o(x^n) = a_0 + P_0(x) + o(x^n)$$

et $g : \mathcal{V}(a_0) \ni V \to E$ définie au voisinage de a_0, admettant un développement limité d'ordre n en a_0 :

$$g(a_0 + h) = b_0 + hb_1 + \cdots + h^n b_n + o(h^n)$$

alors $g \circ f$ admet un développement limité d'ordre n en 0, dont la partie régulière est obtenue en tronquant le polynôme

$$b_0 + P_0(x)b_1 + \cdots + [P_0(x)]^n b_n$$

à l'ordre n.

Démonstration : Comme $\lim_{x \to 0} f(x) = a_0$, la fonction $g \circ f$ est bien définie au voisinage de 0. On pose, pour x voisin de 0, $h(x) = f(x) - a_0$. Comme $h(x)$ **est infiniment petit avec** x, on a

$$g \circ f(x) = g(a_0 + h(x)) = b_0 + h(x)b_1 + \cdots + h(x)^n b_n + o((h(x))^n)$$

Comme $h(x) = a_1 x + \cdots + a_n x^n + o(x^n)$, on a $h(x) = O(x)$ donc $[h(x)]^n = O(x^n)$ et en conséquence

$$g \circ f(x) = b_0 + h(x)b_1 + \cdots + h(x)^n b_n + o(x^n)$$

Enfin, comme h admet un DL_0^n donné par $h(x) = P_0(x) + o(x^n)$, il en est de même des fonctions $(h^i)_{1 \leqslant i \leqslant n}$. Il est alors clair que $g \circ f$ admet un DL_0^n dont la partie régulière est obtenue en tronquant à l'ordre n le polynôme $b_0 + P_0(x)b_1 + \cdots + [P_0(x)]^n b_n$. ∎

Remarquons que cette démonstration donne un moyen de calculer le développement limité de la composée, pour peu que l'on prenne soin de travailler avec des infiniments petits. Le théorème est énoncé en supposant l'existence de développements au même ordre n pour les fonctions f et g et assure l'existence d'un DL_0^n pour $g \circ f$. Si la connaissance d'un développement d'ordre n de f est en général indispensable pour obtenir ce résultat, un examen attentif de la preuve précédente montre que si $a_1 = 0$ et $a_2 \neq 0$, dès que $i > \frac{n}{2}$ la fonction $[h(x)]^i$ est négligeable devant x^n. Il suffira donc de connaître un développement d'ordre égal à la partie entière p de $\frac{n}{2}$ (si le terme complémentaire est par exemple $O(x^{p+1})$ lorsque n est impair) pour obtenir le DL_0^n pour $g \circ f$. Si $a_2 = 0$, $\frac{n}{3}$ serait suffisant.

Par exemple, pour trouver un DL_0^6 de la fonction $\ln(1+\sin^2 x)$, on développera $\ln(1+u)$ à l'ordre 3 et on trouvera (exercice)

$$\ln(1+\sin^2 x) = x^2 - \frac{5}{6}x^4 + \frac{32}{45}x^6 + o\left(x^7\right)$$

Un conseil en tout cas : commencer par développer la fonction f, qui est la première à évaluer quand on calcule $g \circ f$. Cela évite souvent bien des erreurs.

Pour développer $\exp(f(x))$, on se ramènera à développer la fonction exponentielle au voisinage de 0 en écrivant $\exp(f(x)) = e^{a_0}e^{u(x)}$ où u est infiniment petit avec x. De même, si $a_0 > 0$, on développera la fonction $\ln(f(x))$ en l'écrivant $\ln a_0 + \ln(1+u(x))$.

8-6.5.4 Quotient

THÉORÈME 8-6.17 Si f et g sont à valeurs dans $\mathbb{K}$ et possèdent un développement limité à l'ordre n en 0, avec $g(0) \neq 0$ alors $\frac{f}{g}$ admet un développement limité d'ordre n.

Démonstration : En remplaçant g par $\frac{g}{g(0)}$, on peut toujours supposer $g(0) = 1$ et on a alors

$$\frac{f(x)}{g(x)} = f(x) \times \frac{1}{1+h(x)}$$

où $h(x)$ est infiniment petit avec x et admet un développement limité d'ordre n en 0. Comme $u \mapsto \frac{1}{1+u}$ admet le DL_0^n bien connu, le théorème de composition et le résultat relatif au produit montre que $\frac{f}{g}$ admet un développement d'ordre n, qu'on calculera par les techniques décrites précédemment. ■

EXERCICE 8-6.18 Donner le développement limité d'ordre 5 en 0 de $\frac{\sin x}{2+\cos x}$ (réponse $\frac{1}{3}x - \frac{1}{540}x^5 + o\left(x^6\right)$).

REMARQUE 8-6.19 Si g est infiniment petit d'ordre p, on travaillera avec $\frac{x^p f(x)}{g(x)}$ dont on trouvera un développement à l'ordre n

$$\frac{x^p f(x)}{g(x)} = a_0 + a_1 x + \cdots + a_n x^n + o(x^n)$$

(pour peu que l'on connaisse un développement d'ordre $n+p$ pour g!) ce qui donnera le développement généralisé

$$\frac{f(x)}{g(x)} = \frac{a_0}{x^p} + \frac{a_1}{x^{p-1}} + \cdots + a_n x^{n-p} + o(x^{n-p})$$

qui rend les mêmes services qu'un développement limité. Par exemple (exercice)

$$\operatorname{cotan} x = \frac{1}{x} - \frac{1}{3}x - \frac{1}{45}x^3 + o\left(x^4\right)$$

8-6.6 Développements asymptotiques

La relation de domination au voisinage de zéro

$$f \preccurlyeq g \Leftrightarrow f(x) \underset{0}{=} O(g(x))$$

induit une relation d'ordre total sur la famille de fonctions $(f_n)_{n\in\mathbb{N}}$ définies par $f_n(x) = x^n$:

$$f_n \preccurlyeq f_p \Leftrightarrow n \geqslant p$$

La relation d'ordre strict associée est alors la relation de négligeabilité :

$$\forall f, g \in \{f_n,\ n \in \mathbb{N}\} \quad f \prec g \Leftrightarrow (f \preccurlyeq g \text{ et } f \neq g) \Leftrightarrow f(x) \underset{0}{=} o(g(x))$$

Si f est une fonction définie au voisinage de 0 à valeurs complexes (par exemple), écrire un développement limité de f à l'ordre m (on dira aussi *à la précision* f_m), c'est déterminer des entiers

$$n_0 < n_1 < \cdots < n_p \leqslant m$$

c'est-à-dire tels que

$$f_{n_0} \succ f_{n_1} \succ \cdots \succ f_{n_p} \succcurlyeq f_m$$

et des complexes $a_0, \ldots, a_p$ non nuls[7] tels que

$$f = a_0 f_{n_0} + \cdots + a_p f_{n_p} + o(f_m)$$

Par exemple, le développement limité de la fonction $f(x) = \sin x$ à l'ordre 6 en zéro s'écrira

$$f = f_1 - \frac{1}{6} f_3 + \frac{1}{120} f_5 + o(f_6)$$

Les fonctions de référence (on dit aussi de "l'échelle de comparaison") sont ici particulièrement simples : ce sont les fonctions monômes. Il est parfois nécessaire de travailler avec d'autres fonctions de référence. Par exemple, la fonction $g(x) = x^5 \ln x$ définie à droite de zéro admet un développement dans cette échelle de comparaison à la précision x^4, qui est tout simplement

$$g = o(f_4)$$

par contre, elle n'admet pas de développement à la précision f_5. Il serait judicieux, pour étudier des fonctions telles que g, d'élargir la famille de fonctions de référence et de prendre par exemple les fonctions

$$(f_{p,q})_{p,q\in\mathbb{Z}} \text{ définies par } f_{p,q}(x) = x^p (\ln x)^q$$

Ici encore, cette famille (formée de fonctions qui peuvent être infiniment grandes, infiniment petites, et posséder une limite non nulle dans le cas $(p,q) = (0,0)$) est telle que

$$(p,q) \neq (p',q') \Rightarrow f_{p,q} \underset{0}{=} o(f_{p',q'}) \text{ ou } f_{p',q'} \underset{0}{=} o(f_{p,q})$$

[7]On peut avoir aussi uniquement $f = o(f_m)$

La relation d'ordre strict $f \succ g$ n'est pas autre chose sur la famille $(f_{p,q})_{p,q\in\mathbb{Z}}$ que l'ordre :

$$f_{p',q'} \succ f_{p,q} \Leftrightarrow f_{p,q} \underset{0}{=} o(f_{p',q'}) \Leftrightarrow \begin{cases} p' < p \\ \text{ou} \\ p' = p \text{ et } q' > q \end{cases}$$

Par rapport à cette échelle de référence, trouver un développement asymptotique d'une fonction f à la précision $f_{m,n}$, ce sera obtenir des complexes $a_0, \dots, a_k$ non nuls et des fonctions

$$f_{p_0,q_0} \succ f_{p_1,n_1} \succ \cdots \succ f_{p_k,q_k} \succcurlyeq f_{m,n}$$

tels que

$$f = a_0 f_{p_0,q_0} + a_1 f_{p_1,n_1} + \cdots + a_k f_{p_k,q_k} + o(f_{m,n})$$

Il faut se convaincre qu'une telle écriture, si elle existe, est nécessairement unique. Il est aussi extrêmement important d'ordonner les termes de ce développement, en commençant en particulier par le terme prépondérant (celui qui "donne l'équivalent"). Un tel développement se manipulera alors comme un développement limité et rendra des services analogues. Il pourra notamment être tronqué si besoin est.

Les échelles de comparaison en 0 sont souvent $(x^p)_{p\in\mathbb{N}}$, $(x^p)_{p\in\mathbb{Z}}$ ou encore $(x^p(\ln|x|)^q)_{(p,q)\in\mathbb{Z}\times\mathbb{Z}}$. On pourra aussi travailler avec des exposants non entiers. Au voisinage de $+\infty$, les échelles précédentes sont aussi utilisées. On pourra travailler[8] aussi avec $\left(e^{\alpha x}x^\beta(\ln x)^\gamma\right)_{(\alpha,\beta,\gamma)\in\mathbb{R}^3}$. Il s'agit ici de familles stables par multiplication.

EXEMPLE 8-6.20 Donner un développement en $+\infty$ dans l'échelle de comparaison $(x^p(\ln x)^q)_{(p,q)\in\mathbb{Z}\times\mathbb{Z}}$ à la précision $\left(\dfrac{\ln x}{x}\right)^3$ pour la fonction

$$f(x) = (1+x)^{\frac{1}{x}}$$

Il n'est pas absolument évident au départ qu'un tel développement existe. C'est une analyse raisonnable de la fonction f qui amènera au résultat. On commence par écrire la fonction sous forme exponentielle

$$f(x) = e^{\frac{\ln(1+x)}{x}}$$

et l'on pèse d'abord l'exposant

$$\frac{\ln(1+x)}{x} \underset{+\infty}{\sim} \frac{\ln x}{x}$$

[8]Bien entendu, il n'existe pas d'échelle de référence universelle. Les échelles citées ici ne seraient pas d'un grand secours pour étudier, au voisinage de $+\infty$ la fonction

$$f(x) = e^{x^2}\sin x$$

Il est donc naturel d'écrire

$$f(x) = e^{\frac{\ln x}{x}}.e^{\frac{\ln\left(1+\frac{1}{x}\right)}{x}}$$

qui est un produit de deux fonctions qui tendent vers 1 à l'infini. On essaie donc d'obtenir pour chacune de ces fonctions un développement dans l'échelle souhaitée, à la précision $\left(\frac{\ln x}{x}\right)^3$ (si l'une de ces fonctions était équivalente disons à $\frac{(\ln x)^2}{x}$, la seconde devrait être développée à la précision $\frac{\ln x}{x^2}$...). Pour la première c'est simple :

$$e^{\frac{\ln x}{x}} = 1 + \frac{\ln x}{x} + \frac{1}{2}\left(\frac{\ln x}{x}\right)^2 + \frac{1}{6}\left(\frac{\ln x}{x}\right)^3 + o\left(\frac{\ln x}{x}\right)^3$$

Pour le second terme, $\frac{\ln\left(1+\frac{1}{x}\right)}{x} \underset{+\infty}{\sim} \frac{1}{x^2}$, et un DL à l'ordre 1 de la fonction exponentielle est suffisant pour obtenir

$$e^{\frac{\ln\left(1+\frac{1}{x}\right)}{x}} = e^{\left(\frac{1}{x^2}+o\left(\frac{1}{x^2}\right)\right)} = 1 + \frac{1}{x^2} + o\left(\frac{\ln x}{x}\right)^3$$

En multipliant les deux développements asymptotiques, et en ne gardant que les termes significatifs ordonnés, on obtient aisément

$$(1+x)^{\frac{1}{x}} = 1 + \frac{\ln x}{x} + \frac{1}{2}\left(\frac{\ln x}{x}\right)^2 + \frac{1}{x^2} + \frac{1}{6}\frac{\ln^3 x}{x^3} + o\left(\frac{\ln x}{x}\right)^3$$

8-6.7 Applications des développements limités

Rappelons d'abord les règles de savoir-vivre qui régissent l'utilisation des développements limités :

- Utiliser le symbole "=" (et donc ne pas oublier le terme complémentaire en $o(?)$) et pas le symbole "$\sim$".
- Toujours travailler avec des "infiniment petits", ce qui revient souvent à faire une translation pour se ramener au voisinage de 0.
- Prévoir l'ordre auquel on devra faire les développements limités. Par exemple, avec $f_\lambda(x) = \tan x - \sin x + \lambda \,\mathrm{sh}^3 x$, si l'on doit faire une discussion faisant intervenir le comportement de f_λ au voisinage de 0 qui soit valable pour toute valeur du paramètre λ, on utilisera un développement à l'ordre 5 (fonctions impaires), puisqu'on voit bien qu'en général f_λ est infiniment petit d'ordre 3, mais qu'une valeur exceptionnelle de λ peut amener à l'ordre 5 (si les termes en x^5 ne se simplifient pas !). On écrira donc

$$f_\lambda(x) = \left(\frac{1}{2} + \lambda\right)x^3 + \left(\frac{1}{8} + \frac{1}{2}\lambda\right)x^5 + O\left(x^7\right)$$

  ce qui fait évidemment apparaître la valeur exceptionnelle $\lambda = -\frac{1}{2}$.
- Lorsqu'on intègre un développement limité, ne pas oublier la "constante d'intégration". Toute dérivation d'un DL doit se justifier.

8-6.7.1 Levée d'indétermination

Il s'agit de rechercher la partie principale d'un infiniment grand ou infiniment petit. Avant de se lancer dans des calculs longs, il faut garder à l'esprit que faire un DL à l'ordre 1, c'est faire un calcul de dérivée : si f est dérivable en a

$$f'(a) \neq 0 \Rightarrow f(x) - f(a) \underset{a}{\sim} (x-a) f'(a)$$

Par exemple, si on recherche la partie principale pour $x \to a > 0$ de $f(x) = x^a - a^x$, on calcule $f'(x) = ax^{a-1} - a^x \ln a$, ce qui donne $f'(a) = a^a(1 - \ln a)$. On a donc

$$a \neq e \Rightarrow x^a - a^x \underset{a}{\sim} a^a (1 - \ln a)(x - a)$$

Dans le cas particulier $a = e$, un développement à l'ordre 2 s'impose. On pourrait ici calculer $f''(e)$ et appliquer la formule de Taylor-Young, puisque le calcul n'est pas compliqué. On peut aussi écrire

$$f(e+h) = e^e \left(1 + \frac{h}{e}\right)^e - e^e e^h$$

ce qui donne

$$f(e+h) = e^e \left[1 + h + \frac{e-1}{2e} h^2 - 1 - h - \frac{h^2}{2} + o\left(h^2\right)\right]$$

soit

$$x^e - e^x \underset{e}{\sim} -\frac{e^{e-1}}{2}(x-e)^2$$

Plus généralement, le théorème des accroissements finis permet d'obtenir un résultat de "simplification" par une fonction f :

LEMME 8-6.21 Si $f : I \to E$ est de classe C^1, définie au voisinage de a avec $f'(a) \neq 0_E$ et si $u, v : J \to I$ sont définies au voisinage de b avec

$$\lim_{x \to b} u(x) = \lim_{x \to b} v(x) = a$$

alors

$$f(u(x)) - f(v(x)) \underset{b}{\sim} (u(x) - v(x)) f'(a)$$

Démonstration[9] : La fonction f' est continue en a. Si $\varepsilon > 0$ est donné, on peut trouver $\alpha > 0$ tel que

$$\forall t \in [a - \alpha, a + \alpha] \quad \|f'(t) - f'(a)\| \leqslant \varepsilon$$

[9] Avec l'intégrale, on écrirait simplement

$$f(u(x)) - f(v(x)) = \int_{u(x)}^{v(x)} f'(t)\, dt = \int_{u(x)}^{v(x)} f'(a)\, dt + \int_{u(x)}^{v(x)} (f'(t) - f'(a))\, dt$$

et on conclurait aisément. La démonstration que nous donnons ici est valable sous la seule hypothèse de continuité de f' en a. Remarquons aussi que, pour $f'(a) = 0_E$, la conclusion est alors

$$f(u(x)) - f(v(x)) \underset{b}{=} o(u(x) - v(x))$$

La fonction $\varphi : t \mapsto f(t) - tf'(a)$ vérifie $\|\varphi'(t)\| \leqslant \varepsilon$ sur $[a-\alpha, a+\alpha]$ et donc, par inégalité des accroissements finis, on obtient

$$\forall w, z \in [a-\alpha, a+\alpha] \quad \|\varphi(w) - \varphi(z)\| \\ = \|f(w) - f(z) - (w-z) f'(a)\| \leqslant \varepsilon |w-z|$$

Comme par hypothèse $\lim_b u = \lim_b v = a$, on peut trouver $\eta > 0$ tel que, pour tout $x \in [b-\eta, b+\eta]$, on ait $u(x), v(x) \in [a-\alpha, a+\alpha]$. On a alors

$$|x-b| \leqslant \eta \Longrightarrow \\ \|f(u(x)) - f(v(x)) - (u(x) - v(x)) f'(a)\| \leqslant \varepsilon (u(x) - v(x))$$

ce qui donne le résultat. ■

EXEMPLE 8-6.22 Equivalent pour $x \to +\infty$ de

$$f(x) = \arctan \sqrt[3]{\frac{8x^2+x+4}{x^2+x+1}} - \arctan \sqrt[3]{8\cos\frac{1}{x}}$$

On applique le résultat précédent après avoir remarqué que les arguments de la fonction arctan tendent vers 2. On a alors

$$f(x) \underset{+\infty}{\sim} \frac{1}{5}\left(\sqrt[3]{\frac{8x^2+x+4}{x^2+x+1}} - \sqrt[3]{8\cos\frac{1}{x}}\right)$$

On étudie à présent un accroissement de la fonction $t \mapsto \sqrt[3]{t}$ au voisinage de 8. Donc

$$f(x) \underset{+\infty}{\sim} \frac{1}{5} \times \frac{1}{12}\left(\frac{8x^2+x+4}{x^2+x+1} - 8\cos\frac{1}{x}\right)$$

Un développement limité à l'ordre 1 donne immédiatement

$$f(x) \underset{+\infty}{\sim} -\frac{7}{60x}$$

8-6.7.2 Etude locale d'une représentation graphique

Si $f : I \to \mathbb{R}$ est définie au voisinage de a et admet un développement limité de la forme

$$f(x) = a_0 + a_1(x-a) + a_p(x-a)^p + o(x-a)^p \quad \text{avec} \quad p \geqslant 2 \text{ et } a_p \neq 0$$

la représentation graphique de f admet, au point d'abscisse a, une tangente T_a d'équation

$$(T_a) \qquad y = a_0 + a_1(x-a)$$

Le signe de la quantité

$$\Delta(x) = f(x) - (a_0 + a_1(x-a))$$

permet de placer le point $M(x, f(x))$ par rapport à cette tangente : pour $\Delta(x)$ positif, M est situé "au dessus" de T_a. Comme

$$\Delta(x) \underset{a}{\sim} a_p (x-a)^p$$

on connaît le signe de cette quantité *au voisinage de a*. En particulier

- Si p est pair, $\Delta(x)$ ne change pas de signe avec $x-a$. La représentation graphique de f reste *localement* du même côté de T_a. On dit que le point d'abscisse a est, pour le graphe de f, un point d'aspect ordinaire, ou point à concavité.
- Si p est impair, $\Delta(x)$ change de signe avec $x-a$. La représentation graphique de f traverse donc sa tangente T_a lorsque $x-a$ change de signe. On dit que le point d'abscisse a est, pour le graphe de f, un point d'inflexion.

De même, un développement asymptotique au voisinage de $+\infty$ de la forme

$$f(x) = a_1 x + a_0 + \frac{b_p}{x^p} + o\left(\frac{1}{x^p}\right) \quad \text{avec} \quad p \geqslant 1 \text{ et } b_p \neq 0$$

montre que la droite d'équation $y = a_1 x + a_0$ est asymptote[10] (oblique si $a_1 \neq 0$) à la représentation graphique de f pour $x \to +\infty$. De plus, l'équivalent

$$f(x) - a_1 x - a_0 \underset{+\infty}{\sim} \frac{b_p}{x^p}$$

permet de préciser la position de la représentation graphique de f par rapport à cette asymptote *au voisinage de* $+\infty$. Cette étude est en général indispensable pour un tracé précis de la représentation de f. Si le développement asymptotique est valable en $\pm\infty$, la parité de p permet de voir si la position est différente en $+\infty$ et $-\infty$ (cas le plus fréquent où p est impair). Dans le cas d'un développement de la forme

$$f(x) = a_2 x^2 + a_1 x + a_0 + \frac{b_p}{x^p} + o\left(\frac{1}{x^p}\right) \text{ avec } a_2 \neq 0$$

on parlera évidemment de parabole asymptote d'équation $y = a_2 x^2 + a_1 x + a_0$.

EXEMPLE 8-6.23 Existence d'une asymptote pour la représentation de

$$f(x) = (x+3) \arcsin \sqrt{\frac{1+x}{1+2x}}$$

La quantité sous la racine est positive pour tout $x \in \left]-\infty, -1\right] \cup \left]\frac{-1}{2}, +\infty\right[$, et tend vers $\frac{1}{2}$ pour $x \to \pm\infty$. Comme arcsin est définie sur $[-1, 1]$, une étude pour $x \to \pm\infty$ a un sens. Pour obtenir la position de la courbe par rapport à

[10] On a en effet $\lim_{x \to +\infty} f(x) - a_1 x - a_0 = 0$

son asymptote, il est probable qu'un DL à l'ordre 2 de $x \mapsto \arcsin\sqrt{\dfrac{1+x}{2+x}}$ sera suffisant (ne pas oublier le terme en x qui est en facteur devant). On se ramène au voisinage de 0 en posant $x = \dfrac{1}{t}$, et on étudie

$$\varphi(t) = \arcsin\sqrt{\frac{1+t}{2+t}}$$

On dérive cette fonction

$$\varphi'(t) = \frac{1}{2(2+t)\sqrt{(1+t)}} = \frac{1}{4} - \frac{1}{4}t + o(t)$$

On obtient par intégration

$$\varphi(t) = \frac{\pi}{4} + \frac{t}{4} - \frac{t^2}{8} + o(t^2)$$

soit finalement

$$f(x) = (x+3)\left(\frac{\pi}{4} + \frac{1}{4x} - \frac{1}{8x^2} + o(x^{-2})\right) = \frac{\pi}{4}x + \frac{3\pi}{4} + \frac{1}{4} + \frac{5}{8x} + o\left(\frac{1}{x}\right)$$

La droite d'équation $y = \dfrac{\pi}{4}x + \dfrac{3\pi}{4} + \dfrac{1}{4}$ est donc asymptote à la représentation graphique de f. La courbe est située au dessus de l'asymptote au voisinage de $+\infty$, en dessous pour $x \to -\infty$.

8-7 Fonctions convexes

8-7.1 Définition, caractérisation

DÉFINITION 8-7.1 *Soit I un intervalle de $\mathbb{R}$ (non réduit à un point). Une fonction $f : I \to \mathbb{R}$ est dite convexe sur I si et seulement si*

$$\forall a, b \in I \quad \forall t \in [0,1] \quad f((1-t)a + tb) \leqslant t f(a) + (1-t) f(b)$$

Une fonction $g : I \to \mathbb{R}$ est concave sur I si et seulement si $-g$ est convexe, ce qui revient donc à changer le sens de l'inégalité dans la définition.

Une fonction à la fois concave et convexe sur un intervalle I est évidemment une fonction affine. La définition de la convexité[11] d'une fonction f signifie simplement que tous les points du segment $[A, B]$, où A et B sont les points de la représentation graphique de f de coordonnées respectives $(a, f(a))$ et $(b, f(b))$, sont situés "au dessus" de la représentation graphique de f : l'image d'un barycentre de a et b (affectés de coefficients positifs) est inférieure au barycentre des images $f(a)$ et $f(b)$ affectés des mêmes coefficients. Par "associativité" du barycentre, on obtient :

[11]Remarquons que, dans cette définition, a et b jouent des rôles symétriques (ainsi que t et $1-t$), et on pourra donc toujours supposer $a \leqslant b$.

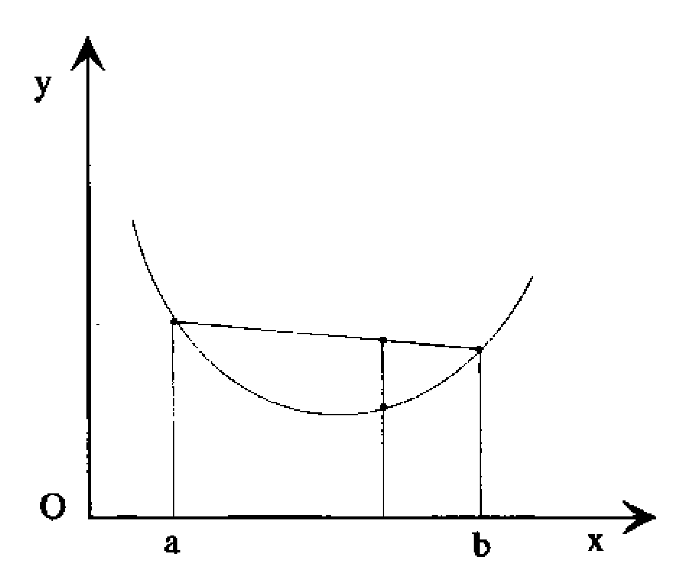


Figure **8.10** – *Fonction convexe*

PROPOSITION 8-7.2 Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si

$$(*) \quad \forall p \in \mathbb{N}^* \quad \forall (a_i)_{1 \leqslant i \leqslant p} \in I \quad \forall (t_i)_{1 \leqslant i \leqslant p} \in [0,1]$$
$$\sum_{i=1}^{p} t_i = 1 \Rightarrow f\left(\sum_{i=1}^{p} t_i a_i\right) \leqslant \sum_{i=1}^{p} t_i f(a_i)$$

Démonstration : Cette condition est évidemment suffisante pour que f soit convexe, puisque avec $p = 2$ on retrouve la définition donnée plus haut. Elle est également nécessaire, puisque pour f est convexe, on prouve $(*)$ par récurrence sur p : l'initialisation pour $p = 1$ ou $p = 2$ ne pose pas de problème. Si on suppose que l'image d'un barycentre (à coefficients positifs) de p points quelconques de I est inférieure au barycentre des images, nous aurons, pour $(a_i)_{1 \leqslant i \leqslant p+1}$ dans I et $(t_i)_{1 \leqslant i \leqslant p+1} \in [0,1]$ de somme égale à 1 (avec $t_{p+1} \in]0,1[$, sinon il n'y a rien à démontrer) :

$$\sum_{i=1}^{p+1} t_i a_i = (1 - t_{p+1}) \left(\sum_{i=1}^{p} \frac{t_i}{1 - t_{p+1}} a_i\right) + t_{p+1} a_{p+1}$$

Par convexité de I, le contenu de la parenthèse appartient à I, et donc

$$f\left(\sum_{i=1}^{p+1} t_i a_i\right) \leqslant (1 - t_{p+1})\, f\left(\sum_{i=1}^{p} \frac{t_i}{1 - t_{p+1}} a_i\right) + t_{p+1} f(a_{p+1})$$

On en déduit la majoration souhaitée au rang $p+1$, puisque l'hypothèse de récurrence donne

$$f\left(\sum_{i=1}^{p} \frac{t_i}{1 - t_{p+1}} a_i\right) \leqslant \sum_{i=1}^{p} \frac{t_i}{1 - t_{p+1}} f(a_i)$$

(puisque $t'_i = \dfrac{t_i}{1 - t_{p+1}} \geqslant 0$ avec $\sum_{i=1}^{p} t'_i = 1$). ■

EXERCICE 8-7.3 Montrer que $f : I \to \mathbb{R}$ est convexe si et seulement si son "épigraphe"

$$\mathcal{E}_f = \{(x, y) \in \mathbb{R}^2 \mid x \in I \quad \text{et} \quad f(x) \leqslant y\}$$

est une partie convexe de $\mathbb{R}^2$.

La caractérisation qui suit est à la base de tous les résultats de continuité et dérivabilité des fonctions convexes :

PROPOSITION 8-7.4 Si I est un intervalle de $\mathbb{R}$, une application $f : I \to \mathbb{R}$ est convexe sur I si et seulement si

$\forall x_0 \in I \quad \varphi_{x_0} : I - \{x_0\} \to \mathbb{R}$ **définie par** $\varphi_{x_0}(t) = \dfrac{f(x) - f(x_0)}{x - x_0}$ **est croissante**

Démonstration : Cette condition est suffisante, puisque si elle est vérifiée on a, pour $a < b \in I$ et $t \in [0, 1]$

$$c = a + t(b - a) \in [a, b]$$

ce qui entraîne

$$\varphi_a(c) \leqslant \varphi_a(b)$$

On en déduit

$$f(c) \leqslant f(a) + \frac{c - a}{b - a}(f(b) - f(a)) = (1 - t) f(a) + f(b)$$

ce qui prouve la convexité de f. Réciproquement, si f est convexe sur I et $x_0 \in I$, montrons que φ_{x_0} est croissante sur $I - \{x_0\}$: si $x_1 < x_2$ sont deux points de $I - \{x_0\}$, on distinguera trois cas suivant que $x_1 < x_2 < x_0$ ou $x_1 < x_0 < x_2$ ou encore $x_0 < x_1 < x_2$. Dans le premier cas,

$$\frac{x_2 - x_1}{x_0 - x_1} = t \in]0, 1[\quad \text{avec} \quad x_2 = (1 - t) x_1 + t x_0$$

La convexité de f montre alors que

$$f(x_2) \leqslant \frac{x_0 - x_2}{x_0 - x_1} f(x_1) + \frac{x_2 - x_1}{x_0 - x_1} f(x_0)$$

ce qui donne

$$f(x_2) - f(x_0) \leqslant \frac{x_2 - x_0}{x_0 - x_1} [f(x_0) - f(x_1)]$$

Comme $x_2 - x_0 < 0$, on déduit

$$\varphi_{x_0}(x_2) \geqslant \varphi_{x_0}(x_1)$$

Les autres cas se traitent de manière analogue, et la croissance de la fonction φ_{x_0} sur $I - \{x_0\}$ en découle. ■

Si M_x désigne le point de la représentation graphique de f ayant x pour abscisse, pour $x \neq x_0$ le nombre $\varphi_{x_0}(x)$ représente la pente de la "sécante" $M_{x_0}M_x$. L'application

$$I - \{x_0\} \to \mathbb{R} \quad x \mapsto \text{pente}(M_{x_0}M_x)$$

est donc croissante. En particulier, si $a < b < c$ sont trois points de I, les inégalités

$$\varphi_a(b) \leqslant \varphi_a(c) = \varphi_c(a) \leqslant \varphi_c(b)$$

s'écrivent aussi

$$\text{pente}(M_aM_b) \leqslant \text{pente}(M_aM_c) \leqslant \text{pente}(M_bM_c)$$

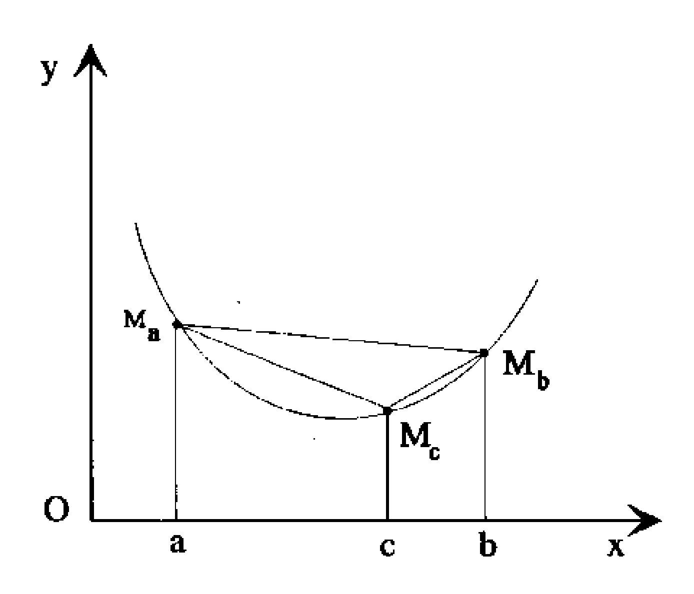


Figure **8.11** – *Inégalités entre pentes*

EXERCICE 8-7.5 Si I et J sont deux intervalles de $\mathbb{R}$, $f : I \to J$ est une fonction convexe et si $g : J \to \mathbb{R}$ est convexe **croissante**, montrer que la composée $g \circ f$ est convexe sur I. En particulier, montrer que toute application $f : I \to \mathbb{R}^{*+}$ telle que $x \mapsto \ln(f(x))$ soit convexe (on dit que f est "Log-convexe") est convexe. Montrer qu'il n'y a pas de réciproque.

EXERCICE 8-7.6 Montrer que, si une suite de fonctions convexes converge simplement sur un intervalle I vers une fonction f, alors f est convexe sur I. Montrer de même que, si $(f_j)_{j\in J}$ est une famille de fonctions convexes sur un intervalle I, et s'il existe $a < b$ dans I avec

$$(f_j(a))_{j\in J} \text{ et } (f_j(b))_{j\in J} \quad \text{majorées}$$

alors la fonction f définie par

$$f(x) = \sup_{j\in J} f_j(x)$$

est finie et convexe sur $[a, b]$. Ce résultat est vrai en particulier si $(f_j)_{j\in J}$ est une famille de fonctions affines.

EXERCICE 8-7.7 Montrer que si f est une fonction convexe sur un segment $[a, b]$, f est majorée et qu'elle atteint sa borne supérieure en a ou en b. Montrer aussi que, si une fonction convexe sur un intervalle présente un minimum local en un point x_0, ce minimum est global.

8-7.2 Continuité et dérivabilité

8-7.2.1 Cas des fonctions dérivables

Nous étudions d'abord le cas particulier des fonctions convexes dérivables sur un intervalle I. La convexité se traduit alors par une croissance de la fonction dérivée :

THÉORÈME 8-7.8 Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction dérivable. Les propriétés suivantes sont équivalentes :

i) **f est convexe sur I.**

ii) **f' est croissante sur I.**

iii) **La représentation graphique de f est située "au dessus" de chacune de ses tangentes.**

Démonstration : Montrons que *i)* $\Rightarrow$ *ii)*. Soient $a < b$ deux points de I. La fonction

$$\varphi : [0,1] \to \mathbb{R} \quad t \mapsto \varphi(t) = f(a + t(b-a)) - (1-t) f(a) - t f(b)$$

est une fonction dérivable (comme composée de fonctions dérivables) négative qui vérifie $\varphi(0) = \varphi(1) = 0$. Comme pour $h \to 0^+$ on a $\varphi(t) = t\varphi'(0) + o(t)$, on a nécessairement $\varphi'(0) \leqslant 0$ et de même $\varphi'(1) \geqslant 0$. Ceci s'écrit

$$f'(a) \leqslant f(b) - f(a) \quad \text{et} \quad f'(b) \geqslant f(b) - f(a)$$

et donc $f'(b) \geqslant f'(a)$: la fonction dérivée f' est donc croissante sur I.

Supposons à présent f' croissante. Pour $x < y$ dans I, le théorème des accroissements finis assure l'existence d'un point $c \in]x, y[$ avec

$$f(y) - f(x) = (y - x) f'(c)$$

La croissance de f' donne alors

$$(y-x) f'(x) \leqslant f(y) - f(x) \leqslant (y-x) f'(y)$$

ce qui peut s'écrire

$$f(y) \geqslant f(x) + (y-x) f'(x) \quad \text{et} \quad f(x) \geqslant f(y) + (y-x) f'(x)$$

Le point M_y de coordonnées $(y, f(y))$ est donc situé au dessus de la tangente à la représentation graphique de f au point d'abscisse x. De même M_x est au dessus de la tangente en M_y : on a donc bien *ii)* $\Rightarrow$ *iii)*.

Supposons enfin que f vérifie la propriété *iii*), et montrons que f est convexe sur I : si $a < b$ sont deux points de I et si $c = (1-t)\,a + tb$ est un point du segment $[a, b]$, les points M_a et M_b sont "au dessus" de la tangente au point M_c. Le segment $[M_a M_b]$ est donc au dessus de cette droite, et donc en particulier

$$(1-t)\,f(a) + tf(b) \geqslant f(c)$$

ce qui prouve la convexité de f. ■

EXERCICE 8-7.9 Montrer que toute fonction convexe dérivable sur un intervalle I est de classe C^1.

Lorsque f est supposée deux fois dérivable sur I, la croissance de f' est donnée par le signe de f'' :

COROLLAIRE 8-7.10 Si f est deux fois dérivable sur un intervalle I, f est convexe si et seulement si f'' est positive sur I.

8-7.2.2 Etude générale

Sur un intervalle ouvert, une fonction convexe possède des propriétés de dérivabilité :

THÉORÈME 8-7.11 Si $f : I \to \mathbb{R}$ est une fonction convexe, f possède en tout point x intérieur à I une demi-dérivée à droite et à gauche vérifiant

$$f'_g(x) \leqslant f'_d(x)$$

Pour $x < y$ intérieurs à I, on a de plus

$$f'_d(x) \leqslant f'_g(y)$$

ces deux inégalités entraînant évidemment la croissance des fonctions f'_d et f'_g sur l'intérieur de I. Si x_0 est un point arbitraire intérieur à I, la représentation graphique de f est située au dessus de chacune de ses demi-tangentes en x_0.

Démonstration : Soient $x < y$ deux points arbitraires intérieurs à I. On peut donc trouver un $h_0 > 0$ tel que

$$[x - h_0, y + h_0] \subset I \quad \text{et} \quad x + h_0 < y - h_0$$

Pour h et $k \in]0, h_0[$, nous avons, d'après les résultats de la section précédente

$$\begin{aligned}\text{pente}\,(M_{x-k}M_x) \leqslant \text{pente}\,(M_x M_{x+h}) &\leqslant \text{pente}\,(M_{x+h}M_{y-k}) \\ &\leqslant \text{pente}\,(M_{y-k}M_y) \leqslant \text{pente}\,(M_y M_{y+h})\end{aligned}$$

Les fonctions croissantes

$$]0, h_0[\ \ni h \mapsto \text{pente}\,(M_x M_{x+h}) \text{ et pente}\,(M_y M_{y+h})$$

sont donc minorées (par pente($M_{x-k}M_x$) et pente($M_{y-k}M_y$) respectivement). Elles présentent donc une limite à droite en 0, ce qui donne l'existence de $f'_d(x)$ et $f'_d(y)$, avec

$$\forall k \in \left]0, h_0\right[\quad \text{pente}\,(M_{x-k}M_x) \leqslant f'_d(x) \leqslant \text{pente}\,(M_{y-k}M_y) \leqslant f'_d(y)$$

Les fonctions décroissantes

$$\left]0, h_0\right[\ni k \mapsto \text{pente}\,(M_{x-k}M_x) \text{ et pente}\,(M_{y-k}M_y)$$

sont à présent majorées (par $f'_d(x)$ et $f'_d(y)$). Elles ont donc une limite en 0, ce qui donne

$$f'_g(x) \leqslant f'_d(x) \leqslant f'_g(y) \leqslant f'_d(y)$$

Si à présent x_0 est un point intérieur à I et $y \in I$ avec $y < x_0$ (le cas $y > x_0$ se traiterait de manière analogue), on a, pour $h > 0$ assez petit

$$\text{pente}\,(M_y M_{x_0}) \leqslant \text{pente}\,(M_{x_0-h}M_{x_0}) \leqslant \text{pente}\,(M_{x_0}M_{x_0+h})$$

ce qui donne, en faisant tendre h vers 0

$$\frac{f(x_0) - f(y)}{x_0 - y} \leqslant f'_g(x_0) \leqslant f'_d(x_0)$$

ce qui donne (puisque $x_0 - y > 0$)

$$f(y) \geqslant f(x_0) + (y - x_0)\, f'_g(x_0) \geqslant f(x_0) + (y - x_0)\, f'_d(x_0)$$

On traduit ainsi le fait que M_y est situé au dessus des demi-tangentes à la représentation graphique de f au point d'abscisse x_0. ■

L'existence d'une dérivée à droite et à gauche assurant la continuité, on a :

COROLLAIRE 8-7.12 Si f est une fonction convexe sur un intervalle ouvert I, alors f est continue sur I.

Par contre, si I contient une de ses extrémités, f n'est pas nécessairement continue en ce point. Un contre-exemple est donnée par la fonction $f : [0,1] \to \mathbb{R}$ définie par $f(x) = 0$ sur $]0,1[$ et $f(\pm 1) = 1$.

EXERCICE 8-7.13 Montrer que toute fonction convexe sur $\mathbb{R}$ majorée est constante.

EXERCICE 8-7.14 Si I est un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{R}$, montrer que f est convexe si et seulement s'il existe une famille $(a_j)_{j \in J}$ de fonctions affines telles que

$$\forall x \in I \quad f(x) = \sup_{j \in J} a_j(x)$$

EXERCICE 8-7.15 Soit I un intervalle ouvert et $f_n : I \to \mathbb{R}$ une suite de fonctions convexes qui converge simplement sur I vers une fonction f. Montrer que la convergence est uniforme sur tout segment.

8-7.3 Inégalités de convexité

La concavité de la fonction logarithme sur $\mathbb{R}^{*+}$ (qui se vérifie à l'aide du signe de la dérivée seconde) donne

$$\forall\,(x_i)_{1\leqslant i\leqslant p}\in\left(\mathbb{R}^{*+}\right)^p\quad\forall\,(m_i)_{1\leqslant i\leqslant p}\in[0,1]^p$$
$$\sum_{i=1}^{p}m_i=1\Rightarrow\ln\left(\sum_{i=1}^{p}m_ix_i\right)\geqslant\sum_{i=1}^{p}m_i\ln(x_i)$$

Par croissance de la fonction exponentielle, ceci peut s'écrire aussi

$$\prod_{i=1}^{p}x_i^{m_i}\leqslant\sum_{i=1}^{p}m_ix_i$$

En particulier, si tous les m_i sont égaux, on obtiendra l'inégalité

$$\forall\,(x_i)_{1\leqslant i\leqslant p}\in\left(\mathbb{R}^{*+}\right)^p\quad\left(\prod_{i=1}^{p}x_i\right)^{\frac{1}{p}}\leqslant\frac{x_1+\cdots+x_p}{p}$$

entre les moyennes géométriques et arithmétiques des $(x_i)_{1\leqslant i\leqslant p}$.

Les inégalités dites de Hölder et Minkowski peuvent aussi se démontrer en utilisant la concavité du logarithme :

DÉFINITION 8-7.16 *Deux réels p et q de $]1,+\infty[$ sont dits exposants conjugués si et seulement si*

$$\frac{1}{p}+\frac{1}{q}=1$$

Si u et v sont deux réels positifs, on a alors

$$uv\leqslant\frac{u^p}{p}+\frac{v^q}{q}$$

Cette inégalité est évidente si u ou v est nul. Dans le cas général, on passe au logarithme en remarquant que

$$\ln(uv)=\frac{1}{p}\ln(u^p)+\frac{1}{q}\ln(v^q)$$

PROPOSITION 8-7.17 Si p et q sont deux exposants conjugués et si $(x_i)_{1\leqslant i\leqslant n}$ et $(y_i)_{1\leqslant i\leqslant n}$ sont des réels positifs

$$\sum_{i=1}^{n}x_iy_i\leqslant\left(\sum_{i=1}^{n}x_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}y_i^q\right)^{\frac{1}{q}}$$

Démonstration : Comme on a, pour tout i,

$$x_iy_i\leqslant\frac{x_i^p}{p}+\frac{y_i^q}{q}$$

on obtient

$$\sum_{i=1}^{p} x_i y_i \leqslant \frac{1}{p}\left(\sum_{i=1}^{n} x_i^p\right) + \frac{1}{q}\left(\sum_{i=1}^{n} y_i^q\right)$$

Ceci donne le résultat souhaité dans le cas particulier où

$$\sum_{i=1}^{n} y_i^q = \sum_{i=1}^{n} x_i^p = 1$$

On peut ensuite se ramener à ce cas particulier en raisonnant par homogénéité, c'est à dire en considérant les familles $(x_i')_{1\leqslant i\leqslant n}$ et $(y_i')_{1\leqslant i\leqslant n}$ définies par

$$x_i' = \frac{x_i}{\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}} \quad \text{et} \quad y_i' = \frac{y_i}{\left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}}}$$

(le cas où l'une des sommes $\left(\sum_{i=1}^{n} y_i^q\right)$ ou $\left(\sum_{i=1}^{n} x_i^p\right)$ est nulle ne pose pas problème, puisque l'inégalité à prouver est alors $0 \leqslant 0$). ■

Par inégalité triangulaire, on obtient, dans le cas de familles de complexes $(u_i)_{1\leqslant i\leqslant n}$ et $(v_i)_{1\leqslant i\leqslant n}$

$$\left|\sum_{i=1}^{n} u_i v_i\right| \leqslant \left(\sum_{i=1}^{n} |u_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |v_i|^q\right)^{\frac{1}{q}}$$

Ce résultat est dénommé inégalité de Hölder (ou inégalité de Schwarz si $p = q = 2$).

EXERCICE 8-7.18 Si $(x_i)_{1\leqslant i\leqslant n}$ et $(y_i)_{1\leqslant i\leqslant n}$ sont des réels positifs, et si p est un réel strictement supérieur à 1, montrer que

$$\left(\sum_{i=1}^{n} (x_i + y_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}}$$

Indication : on écrira

$$(x_i + y_i)^p = x_i (x_i + y_i)^{p-1} + y_i (x_i + y_i)^{p-1}$$

et on majorera les sommes

$$\sum_{i=1}^{n} x_i (x_i + y_i)^{p-1} \quad \text{et} \quad \sum_{i=1}^{n} y_i (x_i + y_i)^{p-1}$$

à l'aide de l'inégalité de Hölder, en considérant l'exposant conjugué

$$q = \frac{p}{p-1}$$

EXERCICE 8-7.19 En déduire que, pour $p \geqslant 1$

$$\|(u_1, \ldots, u_n)\|_p = \left(\sum_{i=1}^{n} |u_i|^p\right)^{\frac{1}{p}}$$

définit une norme sur $\mathbb{K}^n$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Pour $U \in \mathbb{K}^n$, déterminer

$$\lim_{p \to +\infty} \|U\|_p$$

8-8 Suites réelles définies par une itération

Nous rappelons dans cette section quelques résultats utiles pour l'étude de suites réelles $(u_n)_{n \in \mathbb{N}}$ définies par la donnée de leur premier terme u_0 et une relation de récurrence de la forme

$$u_{n+1} = f(u_n)$$

où f est une fonction à valeurs réelles définie sur une partie X de $\mathbb{R}$.

Remarquons d'abord que, pour que cette suite soit définie pour toute donnée initiale $u_0 \in X$, il faut (et il suffit) que X **soit stable par** f, c'est à dire

$$f(X) \subset X$$

Pour l'étude de suites réelles récurrentes, on cherchera donc souvent à déterminer des sous-ensembles Y du domaine de définition de f vérifiant

$$f(Y) \subset Y$$

On est alors assuré que le choix d'une condition initiale $u_0 \in Y$ donne une suite dont toutes les valeurs restent dans Y. La recherche de telles parties Y (qui seront des intervalles pour beaucoup d'exemples élémentaires) se fait généralement par des études de variations de fonctions et des **représentations graphiques**, permettant de visualiser, dans des situations simples, le comportement des suites (monotonie notamment). S'il ne remplace pas une démonstration, le dessin est souvent indispensable pour "deviner" le comportement de la suite, qui se prouve ensuite par des considérations diverses. Citons notamment :

- Si $f : Y \to Y$ est une fonction **continue** et si, pour une donnée $u_0 \in Y$, la suite u_n est convergente **vers un point** l **de** Y, alors

  $$f(l) = l$$

  Cela signifie que les limites éventuelles d'une suite pour laquelle $u_0 \in Y$ sont à chercher parmi les points fixes de $f|_Y$ (c'est à dire les points x de Y vérifiant $f(x) = x$) et les points de $\overline{Y} - Y$ (adhérents à Y)[12].

[12] Rappelons que la détermination *a priori* de la valeur d'une limite n'a jamais été preuve de convergence.

- Si $f : Y \to Y$ est **croissante** (dans la pratique, Y sera un intervalle de $\mathbb{R}$), toute suite correspondant à une condition initiale $u_0 \in Y$ sera **monotone,** le sens de monotonie dépendant des positions respectives de u_0 et u_1. Une récurrence simple montre en effet que

$$u_0 \leqslant u_1 \Rightarrow \forall n \in \mathbb{N} \quad u_n \leqslant u_{n+1}$$
$$u_0 \geqslant u_1 \Rightarrow \forall n \in \mathbb{N} \quad u_n \geqslant u_{n+1}$$

  Prouver la convergence de la suite revient alors à montrer qu'elle est bornée.

- Si $f : Y \to Y$ est **décroissante**, l'application $g = f \circ f : Y \to Y$ (définie parce que Y est stable par f) est croissante et vérifie

$$\forall n \in \mathbb{N} \quad u_{n+2} = g(u_n)$$

  On en déduit la monotonie des deux suites extraites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$, ces deux suites étant de monotonies différentes, puisque par exemple

$$u_{2n} \leqslant u_{2n+2} \Rightarrow f(u_{2n}) \geqslant f(u_{2n+2}) \quad \text{soit} \quad u_{2n+1} \geqslant u_{2n+3}$$

- Enfin, si la fonction définie sur Y par $g(x) = f(x) - x$ a un signe constant, la monotonie de la suite en découle immédiatement, puisque

$$u_{n+1} - u_n = g(u_n)$$

8-8.1 Point fixe attractif, point fixe répulsif

Soit Y un intervalle de $\mathbb{R}$ et $f : Y \to Y$ une application de classe C^1. La continuité de f montre alors que les limites éventuelles **dans** Y de la suite vérifiant

$$u_0 \in Y \quad \text{et} \quad u_{n+1} = f(u_n)$$

sont à chercher parmi les points fixes de f, c'est à dire dans l'ensemble

$$\mathcal{F} = \{x \in Y \mid f(x) = x\}$$

L'hypothèse de dérivabilité faite sur f va permettre de classer ces points fixes :

- Si $\boxed{f(x) = x \text{ avec } |f'(x)| < 1}$, on peut trouver un réel $\alpha > 0$ tel que

$$\forall u_0 \in Y \cap [x - \alpha, x + \alpha]$$
$$\text{la suite définie par } u_0 \text{ et } u_{n+1} = f(u_n) \text{ converge vers } x$$

  En effet, considérons le cas où x est un point intérieur[13] à Y. Par continuité de f', on peut alors trouver un $\alpha > 0$ et un réel $k \in [0, 1[$ tel que

$$\forall t \in [x - \alpha, x + \alpha] \quad |f'(t)| \leqslant k < 1$$

[13]Si x était une extrémité de Y, la stabilité de Y par f entraînerait $f'(x) \geqslant 0$, et l'étude se ferait alors sur $[x, x + \alpha]$ ou $[x - \alpha, x]$ selon que x est extrémité droite ou gauche de Y.

On peut en effet choisir $\varepsilon > 0$ arbitraire et trouver $\alpha > 0$ avec

$$|h| \leqslant \alpha \Rightarrow x + h \in Y \text{ et } |f'(x+h) - f'(x)| \leqslant \varepsilon$$

Le nombre $k = \varepsilon + |f'(x)|$ convient (il est strictement inférieur à 1 si ε est choisi assez petit, d'après l'hypothèse $|f'(x)| < 1$). L'intervalle $[x-\alpha, x+\alpha]$ est alors f-stable, puisque par l'inégalité des accroissements finis

$$|t - x| \leqslant \alpha \Rightarrow |f(t) - x| = |f(t) - f(x)| \leqslant k\,|t - x| \leqslant k\alpha < \alpha$$

Si on choisit u_0 dans ce segment, la suite $(u_n)_{n \in \mathbb{N}}$ correspondante y prend toutes ses valeurs et, par récurrence sur n

$$\forall n \in \mathbb{N} \quad |u_n - x| \leqslant k\,|u_{n-1} - x| \leqslant k^n\,|u_0 - x|$$

et donc

$$u_0 \in [x-\alpha, x+\alpha] \Rightarrow \lim_{n \to +\infty} u_n = x$$

Dans ce cas, on dit que x est un **point fixe attractif pour** f. On peut dans ce cas faire les remarques suivantes :

- Si $f'(x) > 0$, en choisissant α assez petit, on aura f' positive sur $[x-\alpha, x+\alpha]$. L'égalité des accroissements finis montre alors que, lorsque $u_0 \in [x-\alpha, x+\alpha]$

  $$\forall n \in \mathbb{N} \quad (u_{n+1} - x) \text{ et } (u_n - x) \text{ ont même signe}$$

  avec $|u_{n+1} - x| \leqslant |u_n - x|$. La convergence de la suite vers x est donc monotone (voir figure **8.12**).

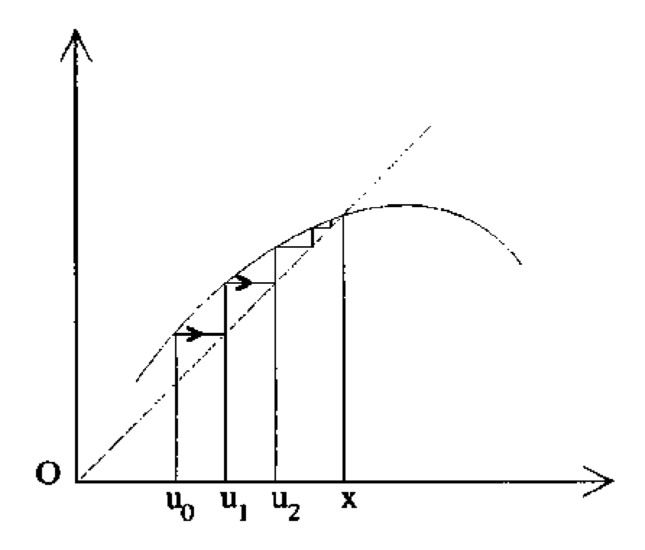


Figure **8.12** – *Convergence monotone vers un point fixe*

– Si $f'(x) < 0$, en choisissant α assez petit, on aura f' négative sur $[x-\alpha, x+\alpha]$. Lorsque l'on choisit $u_0 \in [x-\alpha, x+\alpha]$

$$\forall n \in \mathbb{N} \quad (u_{n+1} - x) \text{ et } (u_n - x) \text{ sont de signes opposés}$$

Les suites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ sont adjacentes, et encadrent x. On visualise cette situation par une "spirale" (figure **8.13**).

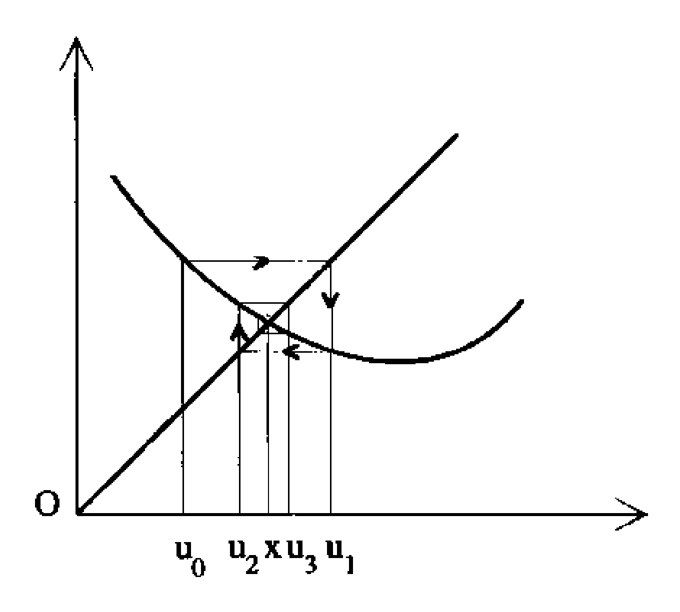


Figure **8.13** – *Convergence d'une itération en "spirale"*

– Lorsque la suite $(u_n)_{n\in\mathbb{N}}$ n'est pas stationnaire et converge vers x, on a $u_n \neq x$ pour tout n et

$$\lim_{n\to+\infty} \frac{|u_{n+1} - x|}{|u_n - x|} = \lim_{n\to+\infty} \left| \frac{f(u_n) - f(x)}{u_n - x} \right| = |f'(x)|$$

et la convergence de la suite sera d'autant plus rapide que $|f'(x)|$ est proche de 0. Plus précisément, si $|f'(x)| < \varepsilon < 1$, on pourra, par continuité, trouver un réel $\beta > 0$ tel que $|f'| \leqslant \varepsilon$ sur $[x-\beta, x+\beta]$. A partir d'un rang p on aura $u_p \in [x-\beta, x+\beta]$, ce qui amènera à la majoration

$$\forall n \geqslant p \quad |u_n - x| \leqslant \varepsilon^{n-p} |u_p - x| = A\varepsilon^n$$

d'autant plus intéressante que ε est petit. La situation "idéale" est $f'(x) = 0$. Dans ce cas, le point x est dit "super-attractif". Nous reviendrons sur ce cas particulier dans la section suivante (méthode de Newton).

- Si $\boxed{f(x) = x \text{ avec } |f'(x)| > 1}$ le point x est dit point fixe répulsif de f : une suite vérifiant $u_{n+1} = f(u_n)$ ne peut converger vers x que si elle est stationnaire. Supposons en effet cette suite définie sur $\mathbb{N}$, avec

$$\forall n \in \mathbb{N} \quad u_n \neq x \quad \text{et} \quad \lim_{n\to+\infty} u_n = x$$

On pourrait alors trouver un réel $k > 1$ et un $\alpha > 0$ tel que

$$\forall t \in [x - \alpha, x + \alpha] \quad |f'(t)| \geqslant k > 1$$

La suite étant convergente vers x sans prendre cette valeur, on aurait

$$\exists p \in \mathbb{N} \quad \forall n \geqslant p \quad u_n \in [x - \alpha, x + \alpha] - \{x\}$$

Le théorème des accroissements finis donnerait alors

$$\forall n \geqslant p \quad |u_{n+1} - x| \geqslant k\,|u_n - x| > |u_n - x|$$

et la suite strictement croissante et positive $(|u_n - x|)_{n \geqslant p}$ ne pourrait alors tendre vers 0.

- Le cas $\boxed{f(x) = x \text{ avec } |f'(x)| = 1}$ correspond à ce qu'on appelle **un point fixe neutre**. Dans ce cas, il peut exister un $\alpha > 0$ tel que tout $u_0 \in [x - \alpha, x + \alpha]$ donne une suite convergente vers x. On peut aussi avoir la situation rencontrée dans le cas d'un point fixe répulsif. Etudier par exemple les cas $Y = \mathbb{R}$ avec $f(x) = \sin x$ et $f(x) = \operatorname{sh} x$.

EXERCICE 8-8.1 Etudier les suites récurrentes définies par u_0,v_0 et $w_0 \in \mathbb{R}^{*+}$ et les relations de récurrence

$$u_{n+1} = \sin(2u_n), \quad v_{n+1} = \frac{1}{2}\left(v_n + \frac{a}{v_n}\right) \quad \text{et} \quad w_{n+1} = \frac{(1 - w_n)^2}{1 + w_n}$$

où a est un réel strictement positif donné.

EXERCICE 8-8.2 On définit l'application $f : \mathbb{R} \to \mathbb{R}$ par

$$f(x) = 4x(1 - x)$$

et on s'intéresse à la suite récurrente associée pour une condition initiale $u_0 \in [0, 1]$. Montrer que $I = [0, 1]$ est un intervalle stable par I. Montrer que, si la suite $(u_n)_{n \in \mathbb{N}}$ est convergente, elle est nécessairement stationnaire. Déterminer les limites éventuelles.

Montrer que, s'il existe un entier $p > 1$ tel que la suite $(u_{np})_{n \in \mathbb{N}}$ converge vers un réel l_0, alors l_0 est un point fixe de

$$f^{(p)} = \underbrace{f \circ \cdots \circ f}_{p \text{ fois}}$$

Montrer qu'alors, pour tout entier $k \in \{0, \ldots, p - 1\}$, la suite $(u_{np+k})_{n \in \mathbb{N}}$ converge vers $l_k = f^{(k)}(l_0)$, qui est aussi un point fixe de $f^{(p)}$.

En utilisant le fait que $f|_{[0,\frac{1}{2}]}$ et $f|_{[\frac{1}{2},1]}$ sont des bijections de leurs intervalles de définition dans $[0, 1]$, donner l'allure des représentations graphiques de $f^{(2)}$, $f^{(3)}$ etc... Montrer que $f^{(p)}$ a exactement 2^p points fixes dans $[0, 1]$.

On se propose de montrer qu'aucun point fixe de $f^{(p)}$ n'est attractif. Pour ce faire, on va utiliser un changement de variable : si $u_0 \in [0, 1]$, on peut trouver un $\theta_0 \in [0, 1[$ avec $u_0 = \sin^2(2\pi\theta_0)$. Montrer qu'alors

$$u_1 = \sin^2(2\pi\theta_1)$$

où le réel $\theta_1 \in [0,1[$ est défini par

$$\theta_1 = 2\theta_0 - [2\theta_0]$$

où $[x]$ désigne la partie entière du réel x. Retrouver les 2^p points fixes de $f^{(p)}$. En utilisant le développement propre de θ_0 en base 2 (*cf.* section **9-2.5**), montrer que, si l est un point fixe de $f^{(p)}$, dans tout intervalle $[l-\alpha, l+\alpha]$ on peut trouver un u_0 pour lequel la suite $(u_{np})_{n\in\mathbb{N}}$ ne converge pas vers l. Conclure.

8-8.2 Application au calcul numérique

Soit g une fonction continue sur un intervalle I de $\mathbb{R}$ telle qu'il existe $a < b \in I$ avec $g(a)g(b) < 0$. Le théorème des valeurs intermédiaires assure l'existence d'au moins[14] un point c de $]a,b[$ vérifiant $g(c) = 0$. On peut procéder par dichotomie pour trouver une suite d'intervalles emboîtés $(I_n = [a_n, b_n])_{n\in\mathbb{N}}$ de longueurs

$$b_n - a_n = \frac{b-a}{2^n}$$

(avec $I_0 = [a_0, b_0] = [a,b]$) contenant un tel zéro de g : il suffit de prendre $c_0 = \frac{1}{2}(a_0 + b_0)$ et, lorsque $g(c_0) \neq 0$, de poser $a_1 = a_0$ et $b_1 = c_0$ si $g(a_0)g(c_0) < 0$, $a_1 = c_0$ et $b_1 = b_0$ dans le cas contraire. Lorsqu'on a un encadrement raisonnable de c, on peut alors utiliser d'autres procédés d'approximation, souvent basés sur des itérations de fonctions : si on peut trouver une fonction f définie au voisinage de c telle que, sur ce voisinage,

$$g(x) = 0 \Leftrightarrow f(x) = x$$

et si de plus f est k-contractante sur un intervalle $[c-\alpha, c+\alpha]$, la suite définie par le choix d'un u_0 "proche de c" et $u_{n+1} = f(u_n)$ convergera vers c, avec (si $u_0 \in [c-\alpha, c+\alpha]$)

$$|u_n - c| \leqslant k^n |u_0 - c|$$

La convergence sera donc plus rapide si k est proche de zéro.

On choisit souvent f de la forme

$$f(x) = x - \frac{g(x)}{\lambda} \quad (*)$$

avec $\lambda \in \mathbb{R}^*$ bien choisi pour que f soit contractante au voisinage de c. Lorsque g est de classe C^1, on prendra λ proche d'une valeur estimée de $g'(c)$, ce qui donnera $f'(c) \approx 0$.

Par exemple, l'équation

$$g(x) = 2\cot x - x = 0$$

possède une unique solution dans l'intervalle $]0, \frac{\pi}{2}[$ (comme le montre un graphique, puis une étude de variation de la fonction). Cette solution n'est sans

[14] L'unicité de c sera assurée si on sait par exemple que g est strictement monotone sur $[a,b]$.

doute pas trop éloignée de la valeur $\frac{\pi}{3} \approx 1.0472$, où la fonction cotan prend la valeur $\frac{\sqrt{3}}{3} \approx 0.5773$. On estime donc

$$g'(c) = -3 - 2\cot^2 c \approx -\frac{11}{3}$$

et on itèrera la fonction

$$f(x) = x + \frac{3}{11} g(x) = \frac{8x + 6\cotan(x)}{11}$$

avec $u_0 = 1$, on obtient les premiers termes de la suite (les calculs sont effectués sur ordinateur avec 20 chiffres significatifs) :

u_0	1.	u_5	1.0768739865123892446
u_1	1.0775050632369076562	u_6	1.0768739863165541385
u_2	1.0768890829079821160	u_7	1.0768739863119161645
u_3	1.0768743439316892119	u_8	1.0768739863118063231
u_4	1.0768739947813842237	u_9	1.0768739863118037217

Une améloration est donnée par **la méthode de Newton**, valable lorsque la dérivée de la fonction g ne s'annule pas au voisinage de c. L'idée est alors de modifier la définition (*) de la fonction f en

$$f(x) = x - \frac{g(x)}{h(x)}$$

où h est une fonction suffisamment régulière ne s'annulant pas au voisinage de c et choisie pour que c soit un point super-attractif de f (on suppose donc au moins g et h de classe C^1). Comme

$$f'(x) = 1 + g(x)\frac{h'(x)}{h^2(x)} - \frac{g'(x)}{h(x)}$$

et puisque $g(c) = 0$, la condition à imposer à h est $g'(c) = h(c)$. Comme c n'est pas connu, on peut prendre pour h une fonction qui vérifie sûrement cette égalité :

$$h = g'$$

On choisira donc

$$f(x) = x - \frac{g(x)}{g'(x)}$$

et la suite donnée par u_0 "proche" de c et

$$u_{n+1} = u_n - \frac{g(u_n)}{g'(u_n)}$$

Cette formule permettant de calculer de proche en proche les termes de la suite a une interprétation géométrique simple : u_{n+1} est l'abscisse du point d'intersection

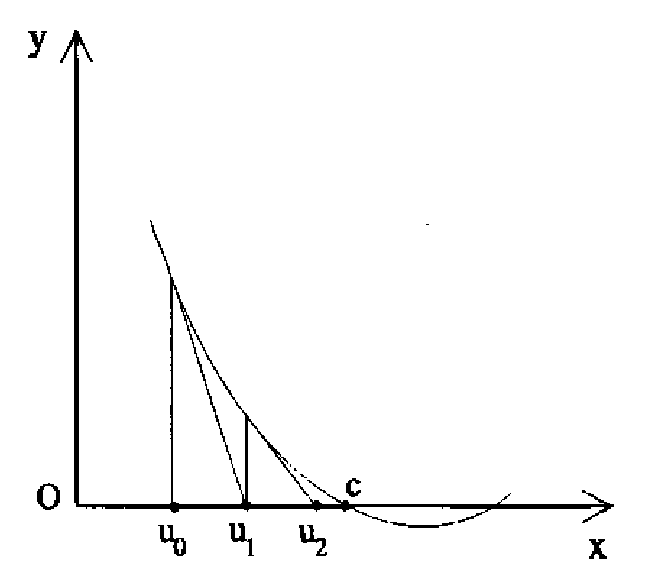


Figure **8.14** – *Méthode de Newton*

de la tangente à la représentation graphique de g au point $M_n(u_n, g(u_n))$ avec l'axe des abscisses.

Si g est de classe C^2 avec g' qui ne s'annule pas au voisinage de c, la fonction f est de classe C^1 au voisinage de c qui en est un point fixe super-attractif. On est donc assuré de l'existence d'un $\alpha > 0$ (pas toujours facile à évaluer en pratique) tel que, pour tout choix de u_0 dans $[c-\alpha, c+\alpha]$, la suite $(u_n)_{n\geqslant 0}$ prenne toutes ses valeurs dans cet intervalle et vérifie

$$\lim_{n\to+\infty} u_n = c$$

On peut mesurer la rapidité de convergence en supposant f de classe C^2 (donc g de classe C^3) : si M est un majorant de f'' sur $[c-\alpha, c+\alpha]$, l'inégalité de Taylor donne

$$u_{n+1} - c = f(u_n) - f(c) = (u_n - c) f'(c) + R_n$$

avec

$$|R_n| \leqslant \frac{1}{2}(u_n - c)^2 M$$

Compte tenu de $f'(c) = 0$, on obtiendra la majoration

$$\forall n \in \mathbb{N} \quad |u_{n+1} - c| \leqslant \frac{M}{2}|u_n - c|^2$$

et on dit alors que la convergence de la suite vers c est **quadratique** : pour tout choix de $u_0 \in [c-\alpha, c+\alpha]$, une récurrence simple donnera

$$\forall n \in \mathbb{N} \quad |u_n - c| \leqslant \frac{2}{M}\left[\frac{M}{2}|u_0 - c|\right]^{2^n}$$

Cette majoration montre que, si par exemple $\frac{M}{2}|u_0 - c| < 10^{-1}$,

$$|u_n - c| \leqslant \frac{2}{M} 10^{-2^n}$$

Chaque itération va alors pratiquement doubler le nombre de décimales exactes dans l'approximation de c par u_n!

Si nous reprenons l'exemple $g(x) = 2\cot x - x$ étudié plus haut, nous aurons alors

$$f(x) = x - \frac{g(x)}{g'(x)} = \frac{2(x + \cot x + x \cot^2 x)}{3 + 2\cot^2 x} = \frac{2x + \sin(2x)}{2 + \sin^2 x}$$

En partant encore de $u_0 = 1$, nous obtenons à présent le tableau

u_0	1.	u_5	1.0768739863118036586
u_1	1.0743052264367312409	u_6	1.0768739863118036586
u_2	1.0768714157682977094	u_7	1.0768739863118036586
u_3	1.0768739863092396755	u_8	1.0768739863118036586
u_4	1.0768739863118036586	u_9	1.0768739863118036586

La convergence ultra-rapide de cette méthode peut être appliquée également à l'extraction de la racine $p^{\text{ième}}$ d'un réel positif a : pour p entier supérieur à 2 et $a > 0$, on prendra

$$g(x) = x^p - a$$

et donc la suite récurrente

$$u_{n+1} = \frac{(p-1)\, u_n^p + a}{p\, u_n^{p-1}}$$

(le cas $p = 2$ correspond à une suite étudiée à l'exercice **8-8**.1) initialisée par un choix "convenable" de u_0.

De même, avec $p = -1$, on obtiendra une suite récurrente

$$u_{n+1} = u_n(2 - a\, u_n)$$

qui permet, pour un choix convenable de u_0 d'approcher $\frac{1}{a}$ en ne faisant que des multiplications. On pourra ainsi "calculer"

$$\frac{b}{a} = b \times \frac{1}{a}$$

pour a et b réels à l'aide de multiplications.

8-9 Exercices

EXERCICE 8-9.1 Soit f croissante de $[x_0, +\infty[$ dans $]0, +\infty[$ telle que

$$\lim_{x\to+\infty} \frac{f(2x)}{f(x)} = 1$$

Etudier pour $c > 0$ $\lim_{x\to+\infty} \frac{f(cx)}{f(x)}$ puis $\lim_{x\to+\infty} \frac{\ln f(x)}{\ln x}$.

EXERCICE 8-9.2 Soit $F : \mathbb{R} \to \mathbb{R}$ deux fois dérivable, avec F et F'' bornées. Montrer que F' est bornée et que $\|F'\|_\infty^2 \leqslant 2\|F\|_\infty \|F''\|_\infty$. (On pourra utiliser la formule de Taylor pour évaluer $F(x \pm h)$). Même question lorsque F est définie sur $\mathbb{R}^+$.Montrer que, si g est C^2 sur $[a, +\infty[$ avec g'' bornée et g tendant vers 0 à l'infini, alors g' tend aussi vers 0 à l'infini.

EXERCICE 8-9.3 Soit $I = [a, b]$, $d \notin [a, b]$ et $f \in C^1[a, b]$ telle que $f(a) = f(b) = 0$. Montrer qu'il existe une tangente au graphe de f passant par $(d, 0)$.

EXERCICE 8-9.4 Soit f une application C^2 de $[a, b]$ dans $\mathbb{R}$ telle que $f(a) = f(b) = 0$. Montrer que :

$$(\forall x \in [a, b]) \quad (\exists c \in]a, b[) \quad f(x) = \frac{(x-a)(x-b)}{2} f''(c)$$

Généraliser.

EXERCICE 8-9.5 Soit P un polynôme de degré impair de $\mathbb{R}[X]$ et $f \in C^\infty(\mathbb{R}, \mathbb{R})$ tels que

$$\forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R}, \quad |f^{(n)}(x)| \leqslant |P(x)|$$

Que peut-on dire de f ?

EXERCICE 8-9.6 Soit $f : \mathbb{R}^{*+} \to \mathbb{R}$ avec, au voisinage de 0 $f(t) = t - at^\alpha + o(t^\alpha)$ où $a > 0$ et $\alpha > 1$.

1. Montrer que, pour t_0 assez petit, la suite $t_{n+1} = f(t_n)$ est définie et converge vers 0.
2. Trouver b et β indépendants de t_0 tels que pour $n \to +\infty$ on ait $t_n \sim bn^\beta$. (On pourra chercher $\gamma \in \mathbb{R}$ tel que $t_n^\gamma - t_{n+1}^\gamma$ ait une limite non nulle).
3. Etudier les exemples $f(x) = \sin(x)$ et $f(x) = \ln(1 + x)$.

EXERCICE 8-9.7 Soient $x_1, \cdots, x_n \in \mathbb{R}^{*+}$. Montrer que

$$\frac{x_1}{x_2} + \frac{x_2}{x_3} + \cdots + \frac{x_{n-1}}{x_n} + \frac{x_n}{x_1} \geqslant n$$

EXERCICE 8-9.8 Soit f une application convexe de $\mathbb{R}^{*+}$ dans $\mathbb{R}$.

1. Montrer que $\frac{f(x)}{x}$ admet une limite l dans $\overline{\mathbb{R}}$ quand x tend vers $+\infty$.
2. Montrer que, si $l \leqslant 0$, f est décroissante.
3. Montrer que, si $l \in \mathbb{R}$, $x \mapsto f(x) - lx$ admet une limite dans $\overline{\mathbb{R}}$ en $+\infty$.

EXERCICE 8-9.9 Déterminer $\lim_{x\to+\infty}\left(\frac{a^{\frac{1}{x}}+b^{\frac{1}{x}}}{2}\right)^x$.

EXERCICE 8-9.10 Soit $f:\mathbb{R}^{*+}\to\mathbb{R}$ telle que $\lim_{x\to 0} f(x)=0$ et

$$\exists\lambda\in]0,1[\quad \exists a\in\mathbb{R}^*\quad \exists\alpha\in\mathbb{R}^{*+}\qquad f(x)-f(\lambda x)\sim ax^\alpha\quad (x\to 0)$$

Montrer que $f(x)\sim\dfrac{ax^\alpha}{1-\lambda^\alpha}\quad (x\to 0)$.

EXERCICE 8-9.11 Pour $n\in\mathbb{N}^*$ trouver les zéros de $f^{(n)}$ avec

$$f(x)=\ln(1+x^2)-\mathrm{Arctan}(x)$$

EXERCICE 8-9.12 Soit $P_n(x)=x^n-nx+1$, et u_n la plus grande racine de P_n.

1. Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ admet une limite l à déterminer.
2. Trouver un développement asymptotique de u_n comportant trois termes.

EXERCICE 8-9.13 Soit $f:\mathbb{R}\to\mathbb{R}$ avec $f(x)=\Big(\mathrm{ch}(x)\Big)^{\frac{1}{x}}$ si $x\neq 0$ et $f(0)=1$.

- Calculer $\lim_{-\infty} f=a$ et $\lim_{+\infty} f=b$.
- Variations et graphe de f.
- Montrer que f admet un développement limité en 0 à tout ordre.
- Montrer que lorsque $x\to+\infty$, f admet un développement asymptotique de la forme

$$f(x)=\sum_{k=0}^{n}\frac{b_k}{x^k}\ +\ o\Big(\frac{1}{x^n}\Big)$$

et que, pour x donné, on n'a pas $f(x)=\sum_{k=0}^{+\infty}\frac{b_k}{x^k}$.

EXERCICE 8-9.14 Soient u et v deux fonctions définies au voisinage de 0 avec

$$\forall x\neq 0\ \ u(x)\neq v(x)$$

On suppose que $\lim_0 u=\lim_0 v=3$. Calculer $\lim_0\dfrac{u^v-v^u}{u-v}$.

EXERCICE 8-9.15 Calculer $\lim_{n\to+\infty}\sum_{k=1}^{n}\sin\left(\frac{1}{k+n}\right)$.

EXERCICE 8-9.16 Soient $1<a_1<\cdots<a_n$ n réels. Quel est le nombre de solutions de l'équation

$$a_1^x+\cdots+a_n^x=A$$

lorsque A est un réel donné? Si x_A est solution, donner un équivalent de x_A pour $A\to+\infty$. Pousser plus loin le développement asymptotique.

EXERCICE 8-9.17 Etudier la convergence simple et uniforme de la suite de fonctions

$$g_n(x) = \cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{2^2}\right)\cdots\cos\left(\frac{x}{2^n}\right)$$

Trouver toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues en 0 et vérifiant

$$\forall x \in \mathbb{R} \quad f(2x) = \exp\left(\frac{x^2}{2}\right)\cos(x)f(x)$$

EXERCICE 8-9.18 Soit $f : \mathbb{R} \to \mathbb{R}$ continue telle que $\forall (x,y) \in \mathbb{R}^2$ $|f(x) - f(y)| \geqslant |x - y|$. Montrer que f est bijective.

EXERCICE 8-9.19 Soit $f : \mathbb{R} \to \mathbb{R}$. On suppose qu'il existe $\alpha > 1$ et $c > 0$ tels que

$$\forall x, y \quad |f(x) - f(y)| \leqslant c|x - y|^\alpha$$

Que peut-on dire de f ?

EXERCICE 8-9.20 Soit $f : \mathbb{R} \to \mathbb{R}$ deux fois dérivable. On suppose qu'il existe $\alpha \in \mathbb{R}$ tel que

$$\lim_{x \to +\infty} f'(x) + \alpha f''(x) = l \in \mathbb{R}$$

Que peut-on dire de $\lim\limits_{x \to +\infty} \dfrac{f(x)}{x}$?

EXERCICE 8-9.21 Soient a_n et b_n deux suite de réels > 0, avec $\lim\limits_{n \to +\infty} a_n^n = a > 0$ et $\lim\limits_{n \to +\infty} b_n^n = b > 0$. Si p et q sont deux réels strictement positifs avec $p + q = 1$, déterminer $\lim\limits_{n \to +\infty} (pa_n + qb_n)^n$.

EXERCICE 8-9.22 Soit f dérivable : $[a, b] \to \mathbb{R}$ avec $f'(a) = f'(b) = 0$. Montrer qu'il existe $c \in]a, b[$ tel que $f'(c) = \dfrac{f(c) - f(a)}{c - a}$. Interprétation géométrique?

EXERCICE 8-9.23 Soit f une fonction à valeurs réelles quatre fois dérivable au voisinage de 0. Si $f''(0) \neq 0$, montrer que l'égalité $f(x) = f(0) + xf'(x\theta(x))$ définit une unique fonction $x \mapsto \theta(x)$ sur un voisinage de 0, à valeurs dans $[0, 1]$. Montrer que θ possède un développement limité d'ordre 2 en 0, dont on donnera la partie régulière.

EXERCICE 8-9.24 Soient $(c_i)_{1 \leqslant i \leqslant p}$ p réels non tous nuls et $(a_i)_{1 \leqslant i \leqslant p}$ p réels tels que $a_1 < \cdots < a_p$. Montrer que l'application $x \longmapsto \sum\limits_{i=1}^{p} c_i \exp(a_i x)$ admet au plus $p-1$ zéros.
Soient $(\alpha_i)_{1 \leqslant i \leqslant n}$ n réels et $(\beta_i)_{1 \leqslant i \leqslant n}$ n réels tels que $\alpha_1 < \cdots < \alpha_n$ et $\beta_1 < \cdots < \beta_n$. Montrer que $\det M > 0$ avec $M = (\exp(\alpha_i \beta_j))_{1 \leqslant i,j \leqslant n}$.

Chapitre 9

Séries dans un espace normé

9-1 Convergence d'une série

Dans tout le chapitre $(\mathbb{E}, \| \ \|)$ est un $\mathbb{K}$-ev normé ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Lorsque $(\mathbb{E}, \| \ \|) = (\mathbb{K}, | \ |)$, les séries étudiées seront dites **séries numériques.**

9-1.1 Série convergente

DÉFINITION 9-1.1 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathbb{E}$. On appelle série de terme général u_n la suite $(S_n)_{n\in\mathbb{N}}$ de $\mathbb{E}^{\mathbb{N}}$ définie par*

$$S_n = \sum_{k=0}^{n} u_k$$

S_n est appelée somme partielle d'ordre n de la série. Lorsque la suite S_n converge dans $(\mathbb{E}, \| \ \|)$, sa limite

$$S = \lim_{n\to+\infty} S_n$$

est appelée somme de la série de terme général u_n, et on écrit symboliquement

$$S = \sum_{k=0}^{+\infty} u_k$$

Attention ! Le symbole $\sum_{k=0}^{+\infty} u_k$ n'a de sens que si l'on sait que la série est convergente. Toute manipulation formelle de ce symbole ne peut être tenue comme valable que si l'on sait (ou l'on admet provisoirement) la convergence de la série. Lorsque la suite des sommes partielles ne converge pas, la série est dite divergente. Nous noterons en abrégé

"$\sum u_n$ CV" pour "la série de terme général u_n est convergente"

"$\sum u_n$ DV" pour "la série de terme général u_n est divergente"

Si le terme général u_n n'est défini qu'à partir d'un rang n_0, on dira, par abus de langage que la série de terme général u_n est convergente si la suite $S_n = \sum_{k=n_0}^{n} u_k$ (définie à partir du rang n_0) possède une limite S, qu'on notera alors $\sum_{k=n_0}^{+\infty} u_k$. Ces notations sont cohérentes, comme le montre la proposition :

PROPOSITION 9-1.2 On ne change pas la nature (convergence ou divergence) de la série $\sum u_n$ en modifiant un nombre fini de termes de la suite u_n. La somme se transforme de manière évidente.

Démonstration : Si $(v_n)_{n\in\mathbb{N}}$ est une suite de $\mathbb{E}$ telle que

$$\forall n > n_0 \quad v_n = u_n$$

on a alors, en notant $U_n = \sum_{k=0}^{n} u_k$ et $V_n = \sum_{k=0}^{n} v_k$

$$\forall n > n_0 \quad U_n - V_n = \sum_{k=0}^{n_0} (u_k - v_k)$$

La différence $U_n - V_n$ étant constante à partir d'un certain rang, la suite U_n est convergente si et seulement s'il en est de même de la suite V_n. Dans le cas de la convergence, on a alors

$$\sum_{k=0}^{+\infty} u_k - \sum_{k=0}^{+\infty} v_k = \sum_{k=0}^{n_0} u_k - \sum_{k=0}^{n_0} v_k \quad \blacksquare$$

Il en résulte que les théorèmes de ce chapitre, énoncés souvent avec une hypothèse sur le terme général u_n supposée vérifiée pour tout entier n, voient leurs conclusions qui subsistent (avec parfois une légère modification) lorsque u_n ne vérifie l'hypothèse qu'à partir d'un rang n_0.

DÉFINITION 9-1.3 *Si $\sum u_n$ est une série convergente, on appelle reste d'ordre n de cette série la différence*

$$R_n = \sum_{k=0}^{+\infty} u_k - \sum_{k=0}^{n} u_k$$

C'est donc

$$R_n = \lim_{m \to +\infty} \sum_{k=0}^{m} u_k - \sum_{k=0}^{n} u_k = \lim_{m \to +\infty} \sum_{k=n+1}^{m} u_k$$

ce qui justifie l'écriture

$$R_n = \sum_{k=n+1}^{+\infty} u_k$$

Cette écriture n'a évidemment de sens que lorsque la série est convergente. Il serait ridicule de parler de reste pour une série divergente.

EXEMPLE 9-1.4 Série géométrique : Si $a \in \mathbb{C}$, la série de terme général a^n est convergente (dans $\mathbb{C}$) si et seulement si $|a| < 1$. On a dans ce cas

$$\sum_{k=0}^{+\infty} a^k = \frac{1}{1-a}$$

Démonstration : Si $a = 1$, on a $\sum_{k=0}^{n} a^k = n+1$, suite qui diverge.

Si $a \neq 1$, on a

$$\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$$

Si $|a| < 1$, on a $\lim_{n \to +\infty} a^n = 0$, et la série est convergente, de somme

$$\sum_{k=0}^{+\infty} a^k = \frac{1}{1-a}$$

Lorsque $|a| \geqslant 1$, on a $|S_n - S_{n-1}| \geqslant 1$, ce qui montre que la suite S_n ne peut converger. Dans le cas où $|a| > 1$, on a d'ailleurs

$$|S_n| \geqslant \frac{|a|^{n+1} - 1}{|a-1|} \underset{n \to +\infty}{\to} +\infty$$

Les opérations sur les suites permettent de définir les opérations analogues sur les séries. On appellera somme des séries de termes généraux u_n et v_n la série de terme général $w_n = u_n + v_n$. Comme

$$\sum_{k=0}^{n} w_k = \sum_{k=0}^{n} u_k + \sum_{k=0}^{n} v_k$$

il est clair que la somme de deux séries convergentes est encore convergente et que, dans ce cas

$$\sum_{k=0}^{+\infty} w_k = \sum_{k=0}^{+\infty} u_k + \sum_{k=0}^{+\infty} v_k$$

On définira de même le produit d'une série par un scalaire. Il est alors clair que

PROPOSITION 9-1.5 $\mathcal{S}(\mathbb{E}) = \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}} \mid \sum u_n \text{ CV}\right\}$ **est un sous-espace vectoriel de $\mathbb{E}^{\mathbb{N}}$, et l'application**

$$\mathcal{S} \to \mathbb{E} \qquad u = (u_n)_{n\in\mathbb{N}} \mapsto \sum_{k=0}^{+\infty} u_k$$

est un morphisme d'espaces vectoriels.

Il est d'ailleurs facile de voir que, si deux des trois séries $\sum u_n$, $\sum v_n$ et $\sum (u_n + v_n)$ convergent, la troisième est aussi convergente. Par contre, la somme de deux séries divergentes peut fort bien être convergente. Il suffit par exemple de prendre $v_n = -u_n$, avec $\sum u_n$ DV.

Lorsque l'espace $\mathbb{E}$ est de dimension finie, on sait que la convergence d'une suite peut se lire coordonnée par coordonnée dans une base de $\mathbb{E}$. En conséquence

PROPOSITION 9-1.6 Si $\dim \mathbb{E}$ **est finie et** $\mathcal{B} = (e_1, \ldots, e_p)$ **est une base de $\mathbb{E}$, pour**

$$u_n = \sum_{i=1}^{p} u_{i,n} e_i \textbf{ avec } (u_{i,n})_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

$\sum u_n$ CV **si et seulement si pour** $i = 1, \ldots, p$ $\quad \sum u_{i,n}$ CV, **et on a alors**

$$\sum_{k=0}^{+\infty} u_k = \sum_{i=1}^{p} \left(\sum_{k=0}^{+\infty} u_{i,k} \right) e_i$$

En particulier, une série complexe converge si et seulement si les séries des parties réelles et imaginaires convergent. L'exemple de la série géométrique montre cependant qu'une série à termes complexes s'étudie souvent sans séparer les parties réelles et imaginaires.

9-1.2 Condition nécessaire de convergence

THÉORÈME 9-1.7 Si $(u_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ **et si la série** $\sum u_n$ **converge, alors**

$$\lim_{n\to+\infty} u_n = 0_{\mathbb{E}}$$

Démonstration : Il suffit de remarquer que $u_n = S_n - S_{n-1}$. ■

Attention ! cette condition n'est pas suffisante pour assurer la convergence de la série !

Le contre-exemple de la suite réelle

$$u_n = \ln\left(1 + \frac{1}{n}\right) \text{ pour } n \geqslant 1$$

donne en effet $\lim_{n\to+\infty} u_n = 0$ mais

$$\sum_{n=1}^{N} u_n = \sum_{n=1}^{N} (\ln(n+1) - \ln n) = \ln(N+1) \underset{N\to+\infty}{\to} +\infty$$

ON DIRA QU'UNE SERIE DIVERGE GROSSIEREMENT SI SON TERME GENERAL NE TEND PAS VERS $0_{\mathbb{E}}$.

9-1.3 Cas d'un espace complet

Lorsque $(\mathbb{E}, \| \|)$ est complet (ce qui est le cas si $\mathbb{E}$ est de dimension finie, en particulier dans le cadre des séries numériques), une condition nécessaire et suffisante pour que la suite des sommes partielles soit convergente est qu'elle vérifie le critère de Cauchy. Soit ici

THÉORÈME 9-1.8 Si $(\mathbb{E}, \| \|)$ est complet, la série de terme général u_n est convergente si et seulement si

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \forall p \in \mathbb{N} \qquad \|u_n + u_{n+1} + \cdots + u_{n+p}\| \leqslant \varepsilon$$

Ce critère sera rarement utilisé tel quel, mais interviendra très souvent de manière cachée (notamment dans l'étude de la convergence absolue). Sa négation est parfois commode pour prouver la divergence d'une série (même si l'espace n'est pas complet), sous la forme suivante :

COROLLAIRE 9-1.9 Si on peut trouver deux fonctions strictement croissantes φ_1 et $\varphi_2 : \mathbb{N} \to \mathbb{N}$ avec $\forall n \quad \varphi_1(n) \leqslant \varphi_2(n)$ telles que

$$\sum_{k=\varphi_1(n)}^{\varphi_2(n)} u_k \text{ ne tende pas vers } 0_{\mathbb{E}} \text{ pour } n \to +\infty$$

alors $\sum u_n$ est divergente.

Démonstration : On remarque en effet que, dans le cas de la convergence

$$\sum_{k=\varphi_1(n)}^{\varphi_2(n)} u_k = S_{\varphi_2(n)} - S_{\varphi_1(n)-1}$$

tend vers $0_{\mathbb{E}}$ comme différence de deux suites ayant la même limite. ∎

EXEMPLE 9-1.10 Divergence de la série harmonique. La série de terme général $u_n = \dfrac{1}{n}$ (pour $n \geqslant 1$) est divergente.

En effet

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}$$

est somme de n termes tous plus grands que $\dfrac{1}{2n}$, ce qui montre que

$$S_{2n} - S_n \geqslant n\frac{1}{2n} = \frac{1}{2}$$

EXERCICE 9-1.11 Existe-t-il une injection $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ telle que $\sum \dfrac{\varphi(n)}{n^2}$ soit convergente ?

9-1.4 Relation suite $\longleftrightarrow$ série

D'un point de vue théorique, il n'y a aucune différence entre la notion de suite et celle de série : étudier la série de terme général u_n, c'est étudier la suite des sommes partielles. Simplement, parce que le terme général s'écrit

$$S_n = u_0 + u_1 + \ldots + u_n$$

nous obtiendrons dans ce chapitre de nombreux résultats qui, à partir d'hypothèses sur u_n, permettront de conclure à la divergence ou à la convergence de la suite S_n.

Ces résultats pourront ensuite être appliqués pour étudier des suites, à l'aide du résultat élémentaire suivant :

PROPOSITION 9-1.12 Si $(v_n)_{n\in\mathbb{N}}$ est une suite de vecteurs de $\mathbb{E}$, on appelle série télescopique associée à cette suite la série de terme général

$$u_n = v_n - v_{n-1}$$

La suite de terme général v_n et la série de terme général u_n sont de même nature (toutes deux convergentes ou divergentes) puisque

$$\sum_{k=1}^{n} u_k = v_n - v_0$$

Dans le cas de la convergence on a

$$\sum_{k=0}^{+\infty} u_k = \lim_{n\to+\infty} v_n - v_0$$

A l'inverse, une série dont le terme général peut s'écrire $u_n = v_n - v_{n-1}$ avec v_n "simple" peut s'étudier facilement, puisque la suite des sommes partielles s'exprime simplement.

EXEMPLE 9-1.13 Convergence et somme de la série de terme général

$$u_n = \frac{1}{n(n+1)(n+2)\cdots(n+p)} \text{ où } p \in \mathbb{N}^* \text{ est fixé et } n \geqslant 1$$

On remarque que

$$u_n = \frac{1}{p}\left[\frac{1}{n(n+1)(n+2)\cdots(n+p-1)} - \frac{1}{(n+1)(n+2)\cdots(n+p)}\right]$$

ce qui donne facilement

$$\sum_{k=1}^{n} u_k = \frac{1}{pp!} - \frac{1}{p(n+1)(n+2)\cdots(n+p)}$$

et prouve la convergence de la série, avec

$$\sum_{k=0}^{+\infty} u_k = \frac{1}{pp!}$$

EXERCICE 9-1.14 Convergence et somme de la série de terme général

$$u_n = \arctan\left(\frac{1}{n^2+n+1}\right) \text{ pour } n \geqslant 1$$

(Indication : $\dfrac{1}{n^2+n+1} = \dfrac{\frac{1}{n} - \frac{1}{n+1}}{1+\frac{1}{n(n+1)}}$).

9-1.5 Regroupements de termes

On s'intéresse ici à des regroupements de termes consécutifs, c'est-à-dire que l'on considère une application

$$\varphi : \mathbb{N} \to \mathbb{N}$$

strictement croissante et vérifiant $\varphi(0) = 0$. Si $(u_n)_{n\in\mathbb{N}}$ est une suite de vecteurs de $\mathbb{E}$, on considère la suite

$$v_n = \sum_{k=\varphi(n)}^{\varphi(n+1)-1} u_k$$

PROPOSITION 9-1.15 Si $\sum u_n$ CV, il en est de même de $\sum v_n$ et

$$\sum_{k=0}^{+\infty} u_k = \sum_{k=0}^{+\infty} v_k$$

> *Démonstration* : Il suffit de remarquer que la suite des sommes partielles de la série de terme général v_n est extraite de celle de la série de terme général u_n, puisque
>
> $$\sum_{k=0}^{n} v_k = \sum_{k=0}^{\varphi(n+1)-1} u_k \quad \blacksquare$$

Bien sûr, il n'y a pas de réciproque, puisque l'exemple de la série divergente de terme général $u_n = (-1)^n$ donne une série convergente si on groupe ses termes deux par deux.

Par contre, si $u_n \underset{n\to+\infty}{\to} 0_{\mathbb{E}}$, on a une réciproque partielle, pour peu que les tailles des paquets restent "raisonnables" :

PROPOSITION 9-1.16 Si $\lim_{n\to+\infty} u_n = 0_{\mathbb{E}}$ et si la suite $(\varphi(n+1) - \varphi(n))_{n\in\mathbb{N}}$ est bornée, la convergence de la série de terme général

$$v_n = \sum_{k=\varphi(n)}^{\varphi(n+1)-1} u_k$$

entraîne celle de $\sum u_n$, vers la même somme, d'après la proposition précédente.

Démonstration : Définissons

$$V = \sum_{k=0}^{+\infty} v_k \text{ et } S_n = \sum_{k=0}^{n} u_k$$

Par définition de la convergence, on peut trouver, pour $\varepsilon > 0$ fixé, un rang N_ε tel que

$$n \geqslant N_\varepsilon \Rightarrow \left\| V - \sum_{k=0}^{n} v_k \right\| = \left\| V - S_{\varphi(n+1)-1} \right\| \leqslant \varepsilon$$

Il existe de plus un rang N'_ε tel que $k \geqslant N'_\varepsilon \Rightarrow \|u_k\| \leqslant \varepsilon$. Posons enfin $M = \sup\left(\varphi(n+1) - \varphi(n)\right)_{n\in\mathbb{N}}$. Soit $n \geqslant \max\left(\varphi(N_\varepsilon), N'_\varepsilon\right)$. Il existe alors un unique entier $p \geqslant N_\varepsilon$ avec

$$\varphi(p) \leqslant n < \varphi(p+1)$$

et on peut écrire

$$\|V - S_n\| \leqslant \left\| V - S_{\varphi(p+1)-1} \right\| + \sum_{k=n}^{\varphi(p+1)-1} \|u_k\|$$

par inégalité triangulaire. Compte tenu du choix de p, on a

$$n \geqslant \max\left(\varphi(N_\varepsilon), N'_\varepsilon\right) \Rightarrow \|V - S_n\| \leqslant \varepsilon + \left(\varphi(p+1) - n\right)\varepsilon \leqslant (M+1)\varepsilon$$

ce qui donne bien la convergence de S_n vers V. ■

Il est à remarquer que, sans l'hypothèse de majoration du nombre de termes dans un paquet, le résultat précédent est en défaut. Prendre par exemple la suite u_n dont les termes seraient

$$(1, -1), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\right), \ldots$$

les parenthèses correspondant aux groupements de termes.

9-2 Séries à termes positifs

On s'intéresse dans cette section aux séries à termes réels positifs (éventuellement à partir d'un certain rang). C'est un cas très important en pratique car, même dans le cas de séries dans un espace vectoriel normé (complet), de nombreuses convergences de séries pourront être prouvées en se ramenant à travailler avec des séries à termes positifs (voir la section sur la convergence absolue).

9-2.1 Résultat de base

THÉORÈME 9-2.1 Si $(u_n)_{n\in\mathbb{N}}$ est une suite de réels positifs, la suite des sommes partielles de la série de terme général u_n est croissante. La série $\sum u_n$ est donc convergente ssi la suite des sommes partielles est majorée. On a alors

$$\sum_{k=0}^{+\infty} u_k = \sup_{n\in\mathbb{N}} \sum_{k=0}^{n} u_k$$

Si la série est à termes positifs à partir d'un rang n_0, on a une conclusion analogue en considérant la suite des sommes partielles à partir du rang n_0. Lorsque la série est divergente, les sommes partielles divergent vers $+\infty$. C'est pourquoi, **pour une série à termes positifs,** la notation $\sum u_n$ CV est parfois remplacée par $\sum u_n < +\infty$. Cette dernière écriture est à proscrire dans le cas de séries réelles dont le terme général n'est pas positif, *a fortiori* dans le cas de série vectorielle !

L'hypothèse $u_n \geqslant 0$ est évidemment indispensable pour appliquer le théorème précédent. Voir le contre-exemple $u_n = (-1)^n$. Bien sûr, tous les résultats qui suivent peuvent être modifiés de façon simple dans le cas de séries à termes négatifs (à partir d'un certain rang). Il suffit dans ce cas de remplacer u_n par $-u_n$.

9-2.2 Comparaison de séries à termes positifs

9-2.2.1 Majoration. Domination. Equivalence

THÉORÈME 9-2.2 Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}} \in (\mathbb{R}^+)^{\mathbb{N}}$. Si, à partir d'un certain rang

$$u_n \leqslant v_n$$

on a

$$\sum v_n \text{ CV} \Rightarrow \sum u_n \text{ CV}$$
$$\sum u_n \text{ DV} \Rightarrow \sum v_n \text{ DV}$$

Démonstration : On peut bien sûr supposer la majoration valable à partir du rang 0. On a alors

$$\forall n \in \mathbb{N} \quad U_n = \sum_{k=0}^{n} u_k \leqslant V_n = \sum_{k=0}^{n} v_k$$

Si la série $\sum v_n$ converge, la suite V_n est majorée. Il en est de même de la suite U_n. Lorsque $\lim_{n\to+\infty} U_n = +\infty$, il en est évidemment de même de la suite V_n. ■

EXEMPLE 9-2.3 Pour $n \geqslant 2$ on a $\frac{1}{n} \leqslant \frac{1}{\sqrt{n}}$ et donc $\sum \frac{1}{\sqrt{n}}$ est divergente. De même

$$\frac{1}{n^2} \leqslant \frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}$$

(série télescopique) et donc $\sum \frac{1}{n^2}$ est convergente.

La quasi-totalité des "critères de convergence" pour les séries à termes positifs ne seront que des réécritures du théorème précédent, avec un vocabulaire différent. Il faudra donc garder à l'esprit que l'hypothèse de termes positifs est fondamentale !

COROLLAIRE 9-2.4 Si u_n et $v_n \geqslant 0$ avec $u_n = O(v_n)$ pour $n \to +\infty$, alors

$$\sum v_n \text{ CV} \Rightarrow \sum u_n \text{ CV}$$

Par contraposée

$$\sum u_n \text{ DV} \Rightarrow \sum v_n \text{ DV}$$

Le résultat précédent est évidemment assuré sous l'hypothèse plus forte de négligeabilité $u_n = o(v_n)$ $(n \to +\infty)$.

COROLLAIRE 9-2.5 Si $u_n \geqslant 0$ à partir d'un certain rang, et s'il existe deux réels a et $b > 0$ tels que

$$\forall n \geqslant n_0 \quad 0 < a \leqslant \frac{u_n}{v_n} \leqslant b$$

alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

Démonstration : On a certainement $v_n \geqslant 0$ à partir d'un certain rang puisque le rapport $\frac{u_n}{v_n}$ est positif. ■

COROLLAIRE 9-2.6 Si $u_n \geqslant 0$ à partir d'un certain rang et si

$$u_n \underset{n \to +\infty}{\sim} v_n$$

alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

Démonstration : L'équivalence entraîne l'existence d'un rang à partir duquel

$$\frac{1}{2} u_n \leqslant v_n \leqslant 2u_n$$

(parce que $u_n \geqslant 0$). ■

Si l'on veut résumer (forcément de manière imprécise) ce qui précède, on peut dire que, plus le terme général (d'une série à termes positifs) tend rapidement vers 0, plus cette série a de chances de converger. Etudier une série à termes positifs, c'est essentiellement "peser" convenablement son terme général. C'est l'idée simple qu'il faut garder à l'esprit, plutôt que d'essayer d'appliquer aveuglément un certain nombre de "règles de convergence".

EXERCICE 9-2.7 Nature de la série de terme général

$$u_n = \frac{n^2}{\ln n + 2^n}$$

EXERCICE 9-2.8 Si $p \in \mathbb{N}$, nature de la série de terme général

$$u_n = \frac{1! + 2! + \cdots + n!}{(n+p)!}$$

EXERCICE 9-2.9 Si $u_n \geqslant 0$, montrer que $\sum u_n$ et $\sum \frac{u_n}{1+u_n}$ sont de même nature.

9-2.2.2 Comparaison logarithmique

Lorsque le terme général u_n se présente sous forme de produits d'un nombre de plus en plus grand de termes, c'est plutôt le rapport $\frac{u_{n+1}}{u_n}$ qu'on étudie, si le calcul de ce rapport s'accompagne de simplifications. On utilise alors le résultat du théorème **8-3**.4 :

THÉORÈME 9-2.10 Si $u_n > 0$ et $v_n > 0$ avec $\frac{u_{n+1}}{u_n} \leqslant \frac{v_{n+1}}{v_n}$ à partir d'un certain rang, alors

$$\sum v_n \text{ CV} \Rightarrow \sum u_n \text{ CV}$$

EXEMPLE 9-2.11 Nature de la série de terme général

$$u_n = \frac{(n+1)(n+2)\cdots 2n}{n^n}$$

On évalue ici

$$\frac{u_{n+1}}{u_n} = \frac{2(n+1)(2n+1)n^n}{(n+1)^{n+1}\,n} = \frac{1}{\left(1+\frac{1}{n}\right)^n}\frac{4n+2}{n} \underset{n\to+\infty}{\to} \frac{4}{e}$$

A partir d'un certain rang $\frac{u_{n+1}}{u_n} \geqslant 1 = \frac{1^{n+1}}{1^n}$, et donc $1 = O(u_n)$ pour $n \to +\infty$. $\sum u_n$ DV donc (car son terme général ne rend pas vers 0).

9-2.2.3 Séries de Riemann

Il s'agit des séries de termes généraux $u_n = \frac{1}{n^\alpha}$, avec $n \geqslant 1$ et $\alpha \in \mathbb{R}^{*+}$. Ces séries serviront souvent de séries de référence pour appliquer les théorèmes de comparaison qui précèdent. On a vu précédemment que cette série diverge pour $\alpha = 1$, donc par minoration diverge pour $\alpha \leqslant 1$. On a vu également précédemment la convergence de $\sum \frac{1}{n^2}$. Par majoration, on en déduit la convergence des séries de Riemann pour $\alpha \geqslant 2$. Il reste donc à traiter le cas $\alpha \in]1, 2[$.

THÉORÈME 9-2.12 La série de terme général

$$u_n = \frac{1}{n^\alpha} \quad \textbf{pour} \quad (n \geqslant 1)$$

est divergente pour $\alpha \leqslant 1$ et convergente pour $\alpha > 1$.

Démonstration : Pour $\alpha > 1$, on a, par inégalité des accroissements finis

$$(\alpha-1)\frac{1}{(n+1)^{\alpha}} \leqslant \frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}} \leqslant (\alpha-1)\frac{1}{n^{\alpha}}$$

ce qui donne facilement

$$\frac{1}{n^{\alpha}} \underset{n\to+\infty}{\sim} \frac{1}{\alpha-1}\left[\frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}}\right]$$

terme général d'une série télescopique convergente, puisque, pour $\alpha > 1$, la suite $\frac{1}{n^{\alpha-1}}$ est convergente. ■

Le résultat précédent permettra de traiter toutes les séries réelles dont le terme général u_n est un infiniment petit ayant un ordre dans l'échelle de comparaison $\left(\frac{1}{n^{\alpha}}\right)_{\alpha\in\mathbb{R}^+}$. Plus généralement, si on peut majorer (à partir d'un certain rang) $u_n \geqslant 0$ par un terme de la forme $\frac{c}{n^{\alpha}}$ avec $c > 0$ constant et $\alpha > 1$, on pourra conclure à la convergence de $\sum u_n$. Inversement, une minoration par $\frac{c}{n^{\alpha}}$ avec $\alpha \leqslant 1$ permettra de conclure à la divergence de $\sum u_n$.

EXERCICE 9-2.13 Nature des séries de termes généraux

$$u_n = \left(1 - \frac{\ln n}{n}\right)^{n} \quad \text{et} \quad v_n = \left(\operatorname{ch}\frac{1}{n}\right)^{-n^{\frac{5}{2}}}$$

EXEMPLE 9-2.14 Séries de Bertrand : discuter suivant les valeurs de α et $\beta \in \mathbb{R}$ la nature de $\sum u_n$ avec

$$u_n = \frac{1}{n^{\alpha}(\ln n)^{\beta}} \quad \text{pour } n \geqslant 2$$

Le comportement à l'infini de u_n est d'abord dicté par la valeur du paramètre α. C'est donc d'abord selon α qu'est menée la discussion. Ensuite, c'est fort probablement la position de α par rapport à 1 qui intervient, car u_n est "proche" de $\frac{1}{n^{\alpha}}$ dans l'échelle de comparaison (à l'infini) $\left(\frac{1}{n^{\lambda}(\ln n)^{\mu}}\right)$. Plus précisément :

- Si $\alpha > 1$, on pense que la série converge. Pour le prouver, on cherche à majorer par une série convergente. On pourra choisir $\gamma \in]1,\alpha[$ et remarquer qu'alors

$$u_n \underset{n\to+\infty}{=} o\left(\frac{1}{n^{\gamma}}\right)$$

- Si $\alpha < 1$, on aura de même, pour $\gamma \in]\alpha, 1[$

$$\frac{1}{n^{\gamma}} \underset{n\to+\infty}{=} o(u_n)$$

ce qui prouve cette fois par minoration la divergence de $\sum u_n$.

- Si $\alpha = 1$, la série est certainement divergente pour $\beta \leqslant 0$, puisqu'alors u_n est minoré au voisinage de l'infini par $\frac{1}{n}$. Si on se limite aux valeurs positives de β, on pourra prouver la convergence de $\sum u_n$ pour $\beta > 1$, et la divergence pour $\beta \leqslant 1$. On pourra par exemple, comme on l'a fait pour les séries de Riemann, chercher un équivalent au voisinage de $+\infty$ de

$$v_n = \frac{1}{(\ln n)^\gamma} - \frac{1}{(\ln (n+1))^\gamma}$$

et ajuster la valeur de γ pour obtenir la nature de la série $\sum u_n$ (utiliser l'inégalité des accroissements finis). Nous reverrons l'étude du cas $\alpha = 1$ dans la section **9-2.3**).

9-2.2.4 Comparaison avec une série géométrique : règle de D'Alembert

Nous disposons à présent des séries de Riemann comme séries de référence. Un autre type de série (encore plus simple) est d'un grand secours : les séries géométriques, puisque pour $a > 0$:

$$0 < a < 1 \Leftrightarrow \sum a^n \text{ CV}$$

Le théorème de comparaison logarithmique donne immédiatement :

THÉORÈME 9-2.15 Si $u_n > 0$ à partir d'un certain rang et si on peut trouver deux réels a et $b > 0$ tels que

$$\forall n \geqslant n_0 \qquad a < \frac{u_{n+1}}{u_n} < b$$

alors

$$b < 1 \Rightarrow \sum u_n \text{ CV}$$
$$a \geqslant 1 \Rightarrow \sum u_n \text{ DV}$$

On peut d'ailleurs dire que, dans le cas $a \geqslant 1$, la suite u_n est croissante à partir d'un certain rang, et que par conséquent la série est alors grossièrement divergente. Dans le cas $a > 1$, on a d'ailleurs $\lim_{n \to +\infty} u_n = +\infty$.

Encadrer une suite "à partir d'un certain rang", c'est souvent calculer une limite. Du théorème précédent découle la **règle de D'Alembert** :

COROLLAIRE 9-2.16 Si $u_n > 0$ à partir d'un certain rang et si

$$\lim_{n \to +\infty} \frac{u_{n+1}}{u_n} = l \textbf{ existe dans } \overline{\mathbb{R}}^+$$

alors

$$l < 1 \Rightarrow \sum u_n \text{ CV}$$
$$l > 1 \Rightarrow \sum u_n \text{ DV}$$

Dans le cas $l = 1$, on ne peut conclure, sauf si la limite 1 est atteinte par valeurs supérieures, auquel cas la série est grossièrement divergente.

Démonstration : Si $l < 1$, on choisira $\varepsilon > 0$ avec $l + \varepsilon < 1$, et on aura, à partir d'un certain rang $\frac{u_{n+1}}{u_n} \leqslant l + \varepsilon$, ce qui permet de conclure à la convergence. Même raisonnement si $l > 1$, où l'on choisira cette fois $\varepsilon > 0$ avec $l - \varepsilon > 1$.

Les exemples $u_n = \frac{1}{n^2}$ et $u_n = \frac{1}{n}$ montrent que $l = 1$ est un cas d'indétermination (sauf si la limite est atteinte par valeurs supérieures). ■

EXERCICE 9-2.17 Discuter selon les valeurs du réel $a > 0$ la nature de la série

$$u_n = \frac{(n+1)(n+2)\cdots 2n}{n^n a^n}$$

(Indication : dans le cas d'indétermination, chercher un équivalent de $\ln\left(\frac{u_{n+1}}{u_n}\right)$ pour conclure à la divergence).

EXERCICE 9-2.18 Règle de Cauchy. Enoncer des résultats analogues à ceux qui précèdent en faisant des hypothèses sur la suite $\sqrt[n]{u_n}$. Discuter en fonction de $x > 0$ la convergence de $\sum u_n$ avec

$$u_n = \frac{(n+1)^n}{n^{n+1}} x^n$$

9-2.2.5 Exercice : Comparaison logarithmique à une série de Riemann

On remarque que, pour $v_n = \frac{1}{n^\gamma}$, on a

$$\frac{v_{n+1}}{v_n} = 1 - \frac{\gamma}{n} + o\left(\frac{1}{n}\right) \text{ pour } n \to +\infty$$

et l'on sait que

$$\sum \frac{1}{n^\gamma} \text{ CV} \Leftrightarrow \gamma > 1$$

En s'inspirant de la démarche suivie pour l'étude des séries de Bertrand, montrer que, si $u_n > 0$ à partir d'un certain rang, et si

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + o\left(\frac{1}{n}\right) \text{ pour } n \to +\infty$$

alors, si $\alpha > 1$, la série $\sum u_n$ converge. Si $\alpha < 1$, la série diverge (c'était évident si $\alpha < 0$).

EXERCICE 9-2.19 Pour $a > 0$, nature de la série de terme général

$$u_n = \frac{n!}{a(a+1)\cdots(a+n)}$$

9-2.3 Comparaison à une intégrale

Rappelons d'abord quelques résultats sur lesquels nous reviendrons dans un chapitre ultérieur d'intégration :

Si $f : [0, +\infty[\to \mathbf{R}^+$ est une fonction continue par morceaux positive, la fonction

$$F(x) = \int_0^x f(t)\,dt$$

est croissante sur $[0, +\infty[$. Si elle est majorée, la fonction f est intégrable sur $\mathbb{R}^+$, et

$$\int_{\mathbb{R}^+} f = \int_0^{+\infty} f(t)\,dt = \lim_{x\to+\infty} F(x)$$

THÉORÈME 9-2.20 Soit $f : [0, +\infty[\to \mathbb{R}^+$ continue par morceaux, positive et décroissante. On pose $u_n = f(n)$. Si f est intégrable sur $\mathbb{R}^+$, la série $\sum u_n$ est convergente. Si f n'est pas intégrable, alors $\sum u_n$ DV.

Démonstration : Remarquons que le résultat persistera si f est positive décroissante sur une demi-droite $[x_0, +\infty[$. La décroissance de la fonction f donne évidemment

$$\forall p \in \mathbb{N} \quad f(p) \geqslant \int_p^{p+1} f(t)\,dt \geqslant f(p+1)$$

En notant $F(x) = \int_0^x f(t)\,dt$ et comme d'habitude $S_n = \sum_{k=0}^{n} u_k$, on obtient facilement

$$\forall n \qquad \int_0^{n+1} f(t)\,dt \leqslant S_n \leqslant f(0) + \int_0^n f(t)\,dt$$

(voir figure **9.1**). Il en résulte aisément que, si F est bornée sur $\mathbb{R}^+$, il en est de même de la suite $(S_n)_{n\in\mathbb{N}}$. De même, si la série $\sum u_n$ CV, on a, pour tout n, $F(n) \leqslant \sum_{k=0}^{+\infty} u_k$, ce qui prouve l'intégrabilité de f, vu la croissance de F. ■

EXEMPLE 9-2.21 Nature de $\sum \frac{1}{n(\ln n)^\beta}$ (série de Bertrand dans le cas "douteux"). On démarre ici avec $n = 2$. On a alors $x \mapsto f(x) = \frac{1}{x(\ln x)^\beta}$ est une fonction décroissante au voisinage de $+\infty$ (c'est clair si $\beta > 0$, seul cas à véritablement étudier, puisque la divergence de la série est évidente lorsque $\beta \leqslant 0$) et pour $x \geqslant 2$

$$\int_2^x \frac{dt}{t(\ln t)^\beta} = \frac{1}{1-\beta}\left[\frac{1}{(\ln t)^{\beta-1}}\right]_2^x$$

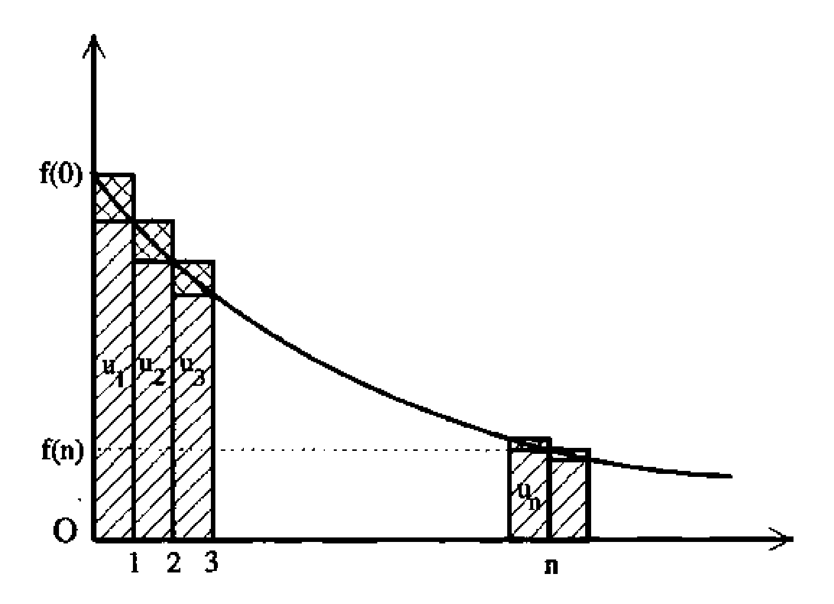


Figure **9.1** – *Comparaison d'une série et d'une intégrale*

si $\beta \neq 1$, alors que

$$\int_2^x \frac{dt}{t(\ln t)} = [\ln(\ln t)]_2^x$$

On s'aperçoit alors que f est intégrable sur $[2, +\infty[$ ssi $\beta > 1$, ce qui donne la condition nécessaire et suffisante pour que $\sum \frac{1}{n(\ln n)^\beta}$ CV.

Une étude un peu plus précise des encadrements obtenus dans la démonstration donne :

COROLLAIRE 9-2.22 Sous les hypothèses du théorème précédent, et dans le cas où $\sum f(n)$ diverge, la suite

$$\Delta_n = \int_0^n f(t)\,dt - \sum_{k=0}^{n} f(k)$$

est monotone bornée, donc convergente. Les "infiniment grands" S_n et $\int_0^n f(t)\,dt$ sont donc "parallèles".

Démonstration : On a

$$\Delta_{n+1} - \Delta_n = \int_n^{n+1} f(t)\,dt - f(n+1) \geqslant 0$$

et donc la suite Δ_n est croissante. On a de plus, d'après les encadrements obtenus précédemment $S_n \geqslant F(n+1) \geqslant F(n)$, et donc $\Delta_n \leqslant 0$. La suite Δ_n est donc convergente. ■

On peut énoncer différemment à la fois le théorème précédent et son corollaire :

THÉORÈME 9-2.23 Si $f : [0, +\infty[\to \mathbb{R}^+$ est continue par morceaux positive et décroissante, la série de terme général (positif)

$$v_n = f(n) - \int_n^{n+1} f(t)\, dt$$

est convergente.

EXEMPLE 9-2.24 **Constante d'Euler :** la fonction $t \mapsto \dfrac{1}{t}$ est décroissante non intégrable sur $[1, +\infty[$. La suite

$$\Delta_n = \int_1^n \frac{dt}{t} - \sum_{k=1}^{n} \frac{1}{k}$$

est donc convergente, et on définit la constante d'Euler par

$$\gamma = \lim_{n \to +\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) - \ln n$$

On démontre que $\gamma \approx 0.5772156649...$

Dans le cas de la convergence, on obtient facilement un encadrement des restes de la série $\sum f(n)$

COROLLAIRE 9-2.25 Lorsque f est positive, intégrable décroissante sur $[a, +\infty[$, on a

$$n \geqslant a \Rightarrow \int_{n+1}^{+\infty} f(t)\, dt \leqslant R_n = \sum_{k=n+1}^{+\infty} f(k) \leqslant \int_n^{+\infty} f(t)\, dt$$

Démonstration : Si $n \geqslant a$, la fonction f étant décroissante sur $[n, +\infty[$

$$\forall k \geqslant n+1 \quad \int_k^{k+1} f(t)\, dt \leqslant f(k) \leqslant \int_{k-1}^{k} f(t)\, dt$$

Il reste ensuite à sommer pour k variant entre $n+1$ et N et à faire tendre N vers $+\infty$. ■

EXEMPLE 9-2.26 Reste de la série de Riemann : pour $\alpha > 1$, montrer que

$$R_n = \sum_{n+1}^{+\infty} \frac{1}{k^\alpha} \underset{n \to +\infty}{\sim} \frac{1}{(\alpha - 1)\, n^{\alpha - 1}}$$

Pour terminer, il faut remarquer que l'idée maîtresse de cette section, qui consiste à "approcher" la somme $\sum_{k=0}^{n} f(k)$ par $\int_0^n f(t)\, dt$ peut être utilisée lorsque f est une fonction croissante (on travaillera encore par encadrements) ou même dans le cas de fonctions vectorielles, où l'on utilisera alors l'intégration par parties lorsqu'elle est possible (voir section **12-5.2**). Donnons ici simplement un exemple :

EXERCICE 9-2.27 Discuter en fonction de α la nature de la série de terme général

$$u_n = \frac{n^\alpha}{(\ln n!)^2}$$

(on cherchera un équivalent de $\ln n!$).

9-2.4 Sommation des relations de comparaison

9-2.4.1 Cas de séries convergentes : restes

THÉORÈME 9-2.28 Soient u_n et $v_n \geqslant 0$ deux suites réelles positives. On suppose que la série $\sum v_n$ CV.

- **Si** $u_n \underset{n\to+\infty}{=} O(v_n)$ **alors** $\displaystyle\sum_{k=n+1}^{+\infty} u_k \underset{n\to+\infty}{=} O\left(\sum_{k=n+1}^{+\infty} v_k\right)$
- **Si** $u_n \underset{n\to+\infty}{=} o(v_n)$ **alors** $\displaystyle\sum_{k=n+1}^{+\infty} u_k \underset{n\to+\infty}{=} o\left(\sum_{k=n+1}^{+\infty} v_k\right)$
- **Si** $u_n \underset{n\to+\infty}{\sim} v_n$ **alors** $\displaystyle\sum_{k=n+1}^{+\infty} u_k \underset{n\to+\infty}{\sim} \sum_{k=n+1}^{+\infty} v_k$

Démonstration : Remarquons d'abord que l'hypothèse de convergence de $\sum v_n$ et les hypothèses de majoration entraînent l'existence des restes des séries, infiniment petits pour $n \to +\infty$ que le théorème permettra donc de comparer. Faisons la démonstration dans le cas de l'équivalence, les autres cas s'étudiant de façon analogue. Si $u_n \underset{n\to+\infty}{\sim} v_n$, on peut trouver, pour $\varepsilon > 0$ fixé, un rang N_ε tel que

$$k > N_\varepsilon \Rightarrow (1-\varepsilon)v_k \leqslant u_k \leqslant (1+\varepsilon)v_k$$

(on notera l'importance de l'hypothèse $v_k \geqslant 0$). On a alors, pour $m > n \geqslant N_\varepsilon$

$$(1-\varepsilon)\sum_{k=n+1}^{m} v_k \leqslant \sum_{k=n+1}^{m} u_k \leqslant (1+\varepsilon)\sum_{k=n+1}^{m} v_k$$

ce qui donne, en faisant tendre m vers $+\infty$

$$n \geqslant N_\varepsilon \Rightarrow (1-\varepsilon)\sum_{k=n+1}^{+\infty} v_k \leqslant \sum_{k=n+1}^{+\infty} u_k \leqslant (1+\varepsilon)\sum_{k=n+1}^{+\infty} v_k$$

ce qui prouve l'équivalence des restes. ■

EXERCICE 9-2.29 Donner des équivalents, pour $n \to +\infty$ de

$$\sum_{k=n}^{+\infty} \frac{1}{k^4+3k-2} \quad \text{et} \quad \sum_{k=n}^{+\infty} \frac{\sqrt{k}}{k!}$$

9-2.4.2 Cas de séries divergentes : sommes partielles

Si $\sum v_n$ DV avec $v_n \geqslant 0$, il n'est plus question de parler de reste pour cette série. Par contre, les sommes partielles

$$V_n = \sum_{k=0}^{n} v_k$$

sont des infiniment grands pour $n \to +\infty$. On obtient des résultats analogues à ceux de la section précédentes :

THÉORÈME 9-2.30 Soient u_n et $v_n \geqslant 0$ deux suites réelles positives. On suppose que la série $\sum v_n$ DV.

- **Si** $u_n \underset{n\to+\infty}{=} O(v_n)$ **alors** $\displaystyle\sum_{k=0}^{n} u_k \underset{n\to+\infty}{=} O\left(\sum_{k=0}^{n} v_k\right)$
- **Si** $u_n \underset{n\to+\infty}{=} o(v_n)$ **alors** $\displaystyle\sum_{k=0}^{n} u_k \underset{n\to+\infty}{=} o\left(\sum_{k=0}^{n} v_k\right)$
- **Si** $u_n \underset{n\to+\infty}{\sim} v_n$ **alors** $\displaystyle\sum_{k=0}^{n} u_k \underset{n\to+\infty}{\sim} \sum_{k=0}^{n} v_k$

Démonstration : Remarquons que, dans le cas de la domination ou de la négligeabilité, on n'a pas d'informations sur la nature de la série $\sum u_n$. Il est à noter que dans ce cas, les conclusions du théorème ne sont significatives que lorsque $\sum u_n$ DV (sinon $\displaystyle\sum_{k=0}^{n} u_k$ possède une limite réelle, et est donc négligeable devant V_n). Faisons la démonstration dans le cas de la négligeabilité : on pose $U_n = \displaystyle\sum_{k=0}^{n} u_k$ et $V_n = \displaystyle\sum_{k=0}^{n} v_k$. Si $\varepsilon > 0$ est fixé, on peut trouver un rang N_ε tel que

$$k \geqslant N_\varepsilon \Rightarrow 0 \leqslant u_k \leqslant \varepsilon v_k$$

Si on choisit $n > N_\varepsilon$, on a alors

$$0 \leqslant U_n = \sum_{k=0}^{N_\varepsilon} u_k + \sum_{k=N_\varepsilon+1}^{n} u_k \leqslant \sum_{k=0}^{N_\varepsilon} u_k + \varepsilon \sum_{k=N_\varepsilon+1}^{n} v_k \leqslant \sum_{k=0}^{N_\varepsilon} u_k + \varepsilon V_n$$

ce qui donne, puisque $V_n > 0$ (au moins à partir d'un certain rang)

$$0 \leqslant \frac{U_n}{V_n} \leqslant \varepsilon + \frac{U_{N_\varepsilon}}{V_n}$$

Comme $\lim\limits_{n\to+\infty} V_n = +\infty$, on peut trouver un rang N' tel que

$$n > N' \Rightarrow \frac{U_{N_\varepsilon}}{V_n} \leqslant \varepsilon$$

On a alors

$$n > \max(N_\varepsilon, N') \Rightarrow 0 \leqslant \frac{U_n}{V_n} \leqslant 2\varepsilon$$

ce qui montre bien que $U_n \underset{n\to+\infty}{=} o(V_n)$. Dans le cas de l'équivalence, on peut écrire $u_n = v_n + w_n$, avec $w_n \underset{n\to+\infty}{=} o(v_n)$. Avec des notations évidentes, on a alors

$$U_n = V_n + W_n$$

avec $|W_n| = \left|\sum\limits_{k=0}^{n} w_k\right| \leqslant \sum\limits_{k=0}^{n} |w_k| \underset{n\to+\infty}{=} o(V_n)$, puisque $|w_n| \underset{n\to+\infty}{=} o(v_n)$.
On a donc $U_n \underset{n\to+\infty}{=} V_n + o(V_n)$, ce qui donne bien

$$U_n \underset{n\to+\infty}{\sim} V_n \quad \blacksquare$$

REMARQUE 9-2.31 Ce résultat contient un résultat classique : le théorème de Cesàro (au moins dans le cas des suites réelles) : si u_n est une suite réelle convergente vers l, alors

$$\lim_{n\to+\infty} \frac{u_1 + u_2 + \cdots + u_n}{n} = l = \lim_{n\to+\infty} u_n$$

Si $l > 0$, on a, à partir d'un certain rang $u_k > 0$. Quitte à modifier un nombre fini de termes de la suite, on peut supposer que $\forall k \quad u_k > 0$ (on ne change pas une limite éventuelle de la suite $\dfrac{u_1 + u_2 + \cdots + u_n}{n}$, puisqu'on lui ajoute une suite de la forme $\dfrac{M}{n}$ avec M constante). On a évidemment $\sum u_n$ DV, puisque le terme général ne tend pas vers 0. Comme $u_n \underset{n\to+\infty}{\sim} l$, on a, d'après le théorème précédent

$$\sum_{k=1}^{n} u_k \underset{n\to+\infty}{\sim} nl$$

ce qui donne le résultat

$$\lim_{n\to+\infty} \frac{\sum\limits_{k=1}^{n} u_k}{n} = l$$

Lorsque $l \leqslant 0$, on peut remplacer u_n par $u_n - l + 1$ par exemple. En fait, l'hypothèse $u_n \in \mathbb{R}$ n'est pas utile en général, et le résultat reste vrai dans tout espace normé (exercice). L'exemple $u_n = (-1)^n$ montre qu'il n'y a pas de réciproque à ce résultat.

EXERCICE 9-2.32 Trouver un équivalent, pour $n \to +\infty$ de

$$S_n = \sum_{k=1}^{n} 2^k \ln(k)$$

(Indication : la fonction ln varie "lentement". On peut deviner que

$$S_n \underset{n\to+\infty}{\sim} \left(\sum_{k=0}^{n} 2^k\right) \ln(n)$$

soit finalement

$$S_n \underset{n\to+\infty}{\sim} 2^{n+1} \ln(n)$$

Pour justifier ce résultat, considérer la série télescopique associée à la suite $2^{n+1}\ln(n)$.)

Les théorèmes précédents sont très utiles pour trouver des développements asymptotiques de suites réelles, puisque le terme général d'une suite peut toujours s'écrire comme somme partielle de la série télescopique associée. Nous verrons aussi dans la section consacrée aux séries absolument convergentes que certains résultats subsistent dans le cas de séries complexes ou vectorielles.

Traitons pour le moment un exemple célèbre :

9-2.4.3 Exemple : formule de Stirling

Il s'agit de l'obtention d'un développement asymptotique de $n!$ pour $n \to +\infty$. Pour des raisons évidentes, on passe d'abord au logarithme :

$$\ln(n!) = \sum_{k=1}^{n} \ln(k) = S_n$$

La croissance de la fonction logarithme, un encadrement par deux intégrales, une intégration par parties

$$\int_1^n \ln t \, dt = n \ln n - n + 1$$

donnent assez facilement

$$S_n \underset{n\to+\infty}{\sim} n \ln n$$

ce qui peut s'écrire aussi $S_n = n \ln n + S'_n$, avec $S'_n \underset{n\to+\infty}{=} o(n \ln n)$. Pour étudier le comportement asymptotique de S'_n, on étudie la série télescopique associée

$$\begin{aligned} u_n &= S'_n - S'_{n-1} \\ &= \ln n - n \ln n + (n-1)\ln(n-1) = (n-1)\ln\left(1 - \frac{1}{n}\right) \underset{n\to+\infty}{\sim} -1 \end{aligned}$$

Cette série est donc (à partir d'un certain rang) à termes négatifs, divergente, et en conséquence

$$S'_n - S'_1 = \sum_{k=2}^{n} \left(S'_k - S'_{k-1}\right) \underset{n\to+\infty}{\sim} -(n-1) \underset{n\to+\infty}{\sim} -n$$

ce qui donne finalement[1]

$$S_n = n\ln n - n + o(n)$$

Pour préciser le terme $o(n)$, on étudie à présent la série télescopique associée à $S_n - \ln n + n$, dont le terme général vaut

$$u_n + 1 = 1 + (n-1)\ln\left(1 - \frac{1}{n}\right) = \frac{1}{2n} + o\left(\frac{1}{n}\right)$$

Ici encore, cette série est à termes positifs (à partir d'un certain rang), divergente et donc

$$S_n - \ln n + n \underset{n\to+\infty}{\sim} \frac{1}{2}\sum_{k=2}^{n}\frac{1}{k} \underset{n\to+\infty}{\sim} \frac{\ln n}{2} \quad \text{(d'après l'exemple } \mathbf{9\text{-}2}.24)$$

On a obtenu

$$S_n = n\ln n - n + \frac{1}{2}\ln n + o(\ln n)$$

Encore un effort ! La série télescopique associée à la suite $S_n - n\ln n + n - \frac{1}{2}\ln n$ a maintenant pour terme général

$$\begin{aligned} 1 + (n-1)\ln\left(1-\frac{1}{n}\right) + \frac{1}{2}\ln\left(1-\frac{1}{n}\right) \\ = 1 + \left(n - \frac{1}{2}\right)\ln\left(1 - \frac{1}{n}\right) = -\frac{1}{12n^2} + o\left(\frac{1}{n^2}\right) \end{aligned}$$

ce qui montre que cette série (à termes négatifs à partir d'un certain rang) est convergente. La suite $S_n - n\ln n + n - \frac{1}{2}\ln n$ est donc convergente, et si on note L sa limite, on a à présent

$$S_n = n\ln n - n + \frac{1}{2}\ln n + L + o(1)$$

Ici, si on veut poursuivre le développement, le point de vue change : l'infiniment petit $o(1)$ est égal à $\delta_n = S_n - n\ln n + n - \frac{1}{2}\ln n -$, et est donc l'opposé du reste d'ordre n de la série télescopique convergente étudiée plus haut. Par application du théorème **9-2**.28 on a

$$\delta_n \underset{n\to+\infty}{\sim} \sum_{k=n+1}^{+\infty}\frac{1}{12k^2} \underset{n\to+\infty}{\sim} \frac{1}{12n} \qquad \text{(d'après l'exemple } \mathbf{9\text{-}2}.26)$$

On obtient

$$S_n = n\ln n - n + \frac{1}{2}\ln n + L + \frac{1}{12n} + o\left(\frac{1}{n}\right)$$

[1] On aurait déjà pu obtenir cette information en étudiant avec précision l'encadrement de S_n par deux intégrales.

EXERCICE 9-2.33 Montrer qu'on peut remplacer ce $o\left(\frac{1}{n}\right)$ par $\frac{K}{n^2}+o\left(\frac{1}{n^2}\right)$ et déterminer K.

En passant à l'exponentielle, on obtient

$$n! = n^n e^{-n}\sqrt{n}e^L e^{\frac{1}{12n}+o\left(\frac{1}{n}\right)} = n^n e^{-n}\sqrt{n}e^L\left(1+\frac{1}{12n}+o\left(\frac{1}{n}\right)\right)$$

On démontrera ultérieurement (voir chapitre d'intégration) que $e^L = \sqrt{2\pi}$, et on obtiendra l'équivalent de Stirling :

$$n! \underset{n\to+\infty}{\sim} n^n e^{-n}\sqrt{2\pi n}$$

9-2.4.4 Remarque

Nous verrons que de nombreux résultats énoncés pour les séries à termes réels de signe constant à partir d'un certain rang persistent dans le cadre des séries absolument convergentes. Il faudrait cependant se garder de croire qu'ils sont toujours valables ! Par exemple, le critère d'équivalence

$$u_n \underset{n\to+\infty}{\sim} v_n \Rightarrow \sum u_n \text{ et } \sum v_n \text{ de même nature}$$

n'est pas utilisable hors du cadre des séries absolument convergentes :

La série de terme général $u_n = \frac{(-1)^n}{n}$ $(n \geqslant 1)$ est convergente puisque sa somme partielle d'indice $2n$

$$S_{2n} = \sum_{k=1}^{n}\left(\frac{1}{2k}-\frac{1}{2k-1}\right) = -\sum_{k=1}^{n}\frac{1}{2k(2k-1)}$$

est somme partielle d'une série à termes négatifs convergente (terme général équivalent à $-\frac{1}{4k^2}$), et que $S_{2n+1} = S_{2n} - \frac{1}{2n+1}$ converge vers la même limite[2]. On a avec $v_n = u_n + \frac{1}{n\ln n}$

$$u_n \underset{n\to+\infty}{\sim} v_n, \quad \sum u_n \text{ CV et } \sum v_n \text{ DV}$$

car $\sum \frac{1}{n\ln n}$ est divergente.

[2]En écrivant

$$S_{2n} = \frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2n}-\left(1+\frac{1}{3}+\cdots+\frac{1}{2n-1}\right)$$

on obtient facilement

$$S_{2n} = \sum_{k=1}^{n}\frac{1}{k}-\sum_{k=1}^{2n}\frac{1}{k}$$

ce qui, en faisant intervenir la constante d'Euler γ peut s'écrire

$$S_{2n} = \left(\ln n+\gamma+o\left(\frac{1}{n}\right)\right)-\left(\ln 2n+\gamma+o\left(\frac{1}{n}\right)\right)$$

et montre que la somme de cette série vaut $-\ln 2$

9-2.5 Développement décimal d'un réel

Soit x un réel positif. Si n est la partie entière de x (dont l'existence est, rappelons-le, conséquence de la propriété de la borne supérieure), on a

$$x = n + y$$

avec $y = x - n \in [0, 1[$.

Lorsque $n \neq 0$, nous pouvons écrire cette partie entière en base 10 :

$$\exists! \, p \in \mathbb{N} \quad 10^p \leqslant n < 10^{p+1}$$

et nous savons qu'il existe un unique $(\alpha_0, \alpha_1, \dots, \alpha_p) \in \{0, \dots, 9\}^{p+1}$ tel que

$$\alpha_p \neq 0 \text{ et } n = \sum_{k=0}^{p} \alpha_k \, 10^k$$

ce qui s'écrit en numération décimale

$$n = \alpha_p \alpha_{p-1} \cdots \alpha_0$$

Nous nous intéressons à présent au problème de l'écriture décimale de y, c'est à dire à ce qui "viendra après la virgule" dans l'écriture de x.

9-2.5.1 Développement décimal propre

Si $y \in [0, 1[$ et $n \in \mathbb{N}^*$, notons $p_n = [10^n y]$ la partie entière de $10^n y$. Nous avons donc

$$p_n \leqslant 10^n y < p_n + 1$$

et donc

$$\frac{p_n}{10^n} \leqslant y < \frac{p_n}{10^n} + \frac{1}{10^n}$$

Le décimal

$$y_n = \frac{p_n}{10^n}$$

est donc valeur approchée de y par défaut à 10^{-n} près, et on a évidemment

$$\lim_{n \to +\infty} y_n = y$$

Comme $10 \, y \in [0, 10[$, nous avons $p_1 \in \{0, \dots, 9\}$. Etudions à présent ce qui différencie y_{n+1} de y_n : comme $p_n \leqslant 10^n y < p_n + 1$, nous avons

$$10 \, p_n \leqslant 10^{n+1} y < 10 \, p_n + 10$$

et donc

$$10 \, p_n \leqslant p_{n+1} \leqslant 10 \, p_n + 9$$

soit encore

$$a_{n+1} = p_{n+1} - 10\,p_n \in \{0, \dots, 9\} \quad \text{avec} \quad y_{n+1} - y_n = \frac{a_{n+1}}{10^{n+1}}$$

En posant $a_1 = p_1$, nous aurons

$$y_n = y_1 + \sum_{k=1}^{n-1} (y_{k+1} - y_k) = \sum_{k=1}^{n} \frac{a_k}{10^k}$$

et la convergence de la suite $(y_n)_n$ vers y peut se traduire par la convergence de la série de terme général

$$u_n = \frac{a_n}{10^n}$$

avec

$$y = \sum_{n=1}^{+\infty} \frac{a_n}{10^n}$$

ce que l'on notera aussi

$$y = 0, a_1 a_2 \cdots a_{n-1} a_n a_{n+1} \cdots$$

DÉFINITION 9-2.34 *La représentation du réel $y \in [0, 1[$ obtenue par la méthode précédente*

$$y = \sum_{n=1}^{+\infty} \frac{a_n}{10^n} \quad \textit{avec} \quad \forall n \in \mathbb{N}^* \quad a_n \in \{0, \dots, 9\}$$

est appelée développement décimal propre (illimité) de y.

Si $x = n + y$ est le réel positif considéré au début de cette section, la représentation décimale propre de x sera donc

$$x = \alpha_p \alpha_{p-1} \cdots \alpha_0, a_1 a_2 \cdots a_{n-1} a_n a_{n+1} \cdots$$

REMARQUE 9-2.35 Ce que nous avons fait en base 10 se généralise évidemment à une base quelconque : si $b \geqslant 2$ est un entier naturel, tout réel de $[0, 1[$ peut s'écrire

$$y = \sum_{n=1}^{+\infty} \frac{a_n}{b^n} \quad \text{avec } \forall n \quad a_n \in \{0, \dots, b-1\}$$

Le cas $b = 2$ (représentation binaire) est très important en pratique.

9-2.5.2 Unicité du développement

Soit y un réel de $[0,1[$. Un développement décimal de y est une écriture de ce nombre sous la forme

$$(*) \quad y = \sum_{n=1}^{+\infty} \frac{b_n}{10^n} \quad \text{avec } \forall n \quad b_n \in \{0, \dots, 9\}$$

(la majoration $0 \leqslant b_n 10^{-n} \leqslant 9\,10^{-n}$ prouve qu'une série de cette forme est toujours convergente). Nous savons que y possède au moins une écriture de cette forme, à savoir son développement décimal propre. Nous étudions ici l'unicité d'une telle représentation de y.

Supposons que y vérifie l'égalité $(*)$. On a alors

$$\forall n \in \mathbb{N}^* \quad 10^n\, y = \sum_{k=1}^{n} b_k\, 10^{n-k} + \sum_{k=1}^{+\infty} \frac{b_{n+k}}{10^k}$$

La première somme est un nombre entier. La seconde est inférieure à

$$9 \sum_{k=1}^{+\infty} \frac{1}{10^k} = \frac{9}{10} \frac{1}{1 - \frac{1}{10}} = 1$$

et il n'y a évidemment égalité que si la suite $(b_k)_{k \geqslant n+1}$ est constante égale à 9. On a donc deux possibilités :

- **La suite $(b_n)_{n \geqslant 1}$ prend une infinité de fois une valeur différente de** 9 ;

  Le calcul précédent montre alors que

  $$\forall n \in \mathbb{N}^* \quad [10^n y] = \sum_{k=1}^{n} b_k\, 10^{n-k}$$

  et donc

  $$\forall n \in \mathbb{N}^* \quad y_n = \frac{[10^n y]}{10^n} = \sum_{k=1}^{n} \frac{b_k}{10^k}$$

  soit encore

  $$b_1 = [10\, y] \quad \text{et} \quad \forall n \in \mathbb{N}^* \quad \frac{b_{n+1}}{10^{n+1}} = y_{n+1} - y_n$$

  Le développement décimal de y considéré est donc le développement décimal propre.

- **La suite $(b_n)_{n \geqslant 1}$ est stationnaire égale à 9 à partir d'un rang que nous noterons** $p+1$ (avec $p \geqslant 1$, car $p = 0$ donnerait $y = 1$) :

  On a alors $b_p \leqslant 8$ et

  $$10^p\, y = \sum_{k=1}^{p} b_k\, 10^{p-k} + \sum_{k=1}^{+\infty} \frac{9}{10^k} = \sum_{k=1}^{p} b_k\, 10^{p-k} + 1$$

ce qui montre que y est un nombre décimal, dont le développement décimal propre est

$$y = 0, a_1 a_2 \ldots a_p 0000 \ldots$$

avec $a_k = b_k$ pour $k < p$ et $a_p = b_p + 1$.

Résumons l'étude précédente :

PROPOSITION 9-2.36 Tout réel de $[0, 1[$ qui n'est pas décimal possède un unique développement décimal illimité, qui est son développement propre. Un décimal de $]0, 1[$ possède deux développements décimaux : un développement propre qui se termine par une suite de 0 et un développement impropre qui se termine par une suite de 9.

REMARQUE 9-2.37 En base $b \geqslant 2$, ce sont les rationnels de la forme $a\, b^{-n}$ avec $a \in \mathbb{N}^*$ qui admettent deux développements illimités en base b : l'un se termine par une suite de 0, l'autre par une suite de symboles représentant le "chiffre" $b - 1$.

9-2.5.3 Développement décimal d'un rationnel

Un développement décimal d'un réel de $[0, 1[$ est dit périodique s'il existe des entiers $p \geqslant 0$ et $q \geqslant 1$ tels que ce développement s'écrive

$$0, a_1 \ldots a_p a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} \ldots$$

La suite des décimales est donc périodique de période q à partir du rang $p + 1$.

PROPOSITION 9-2.38 Un réel de $[0, 1[$ est rationnel si et seulement s'il admet un développement décimal périodique.

Démonstration : Remarquons que, dans le cas où y est décimal (donc rationnel), ses deux développements décimaux sont périodiques.

Supposons que $y \in [0, 1[$ admette le développement décimal périodique

$$y = 0, a_1 \ldots a_p a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} \ldots$$

Notons m_1 et m_2 les entiers naturels admettant l'écriture décimale

$$m_1 = a_1 \ldots a_p \quad \text{et} \quad m_2 = a_{p+1} \ldots a_{p+q}$$

(avec $m_1 = 0$ si $p = 0$). Nous avons

$$10^p\, y = m_1 + 0, a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} a_{p+1} \ldots a_{p+q} \ldots$$

ce qui donne

$$10^q\, (10^p y - m_1) = m_2 + (10^p y - m_1)$$

soit

$$y = \frac{m_2 + m_1 (10^q - 1)}{10^p (10^q - 1)} \in \mathbb{Q}$$

Réciproquement, si y est un rationnel écrit sous forme irréductible

$$y = \frac{p}{q}$$

avec $p \wedge q = 1$, et si

$$y = 0, a_1 a_2 \ldots a_n \ldots$$

est le développement décimal propre de y, nous avons

$$r_n = 0, a_{n+1} a_{n+2} \ldots = 10^n y - [10^n y] = \frac{10^n p}{q} - p_n = \frac{10^n p - p_n q}{q} < 1$$

C'est un rationnel de numérateur compris entre 0 et $q-1$, qui ne peut donc prendre qu'un nombre fini de valeurs. On peut donc trouver $n_0 < n_1$ avec

$$r_{n_0} = r_{n_1}$$

Par unicité du développement décimal propre, on obtient

$$\forall p \geqslant 1 \quad a_{n_0+p} = a_{n_1+p}$$

ce qui montre que la suite $(a_n)_{n \in \mathbb{N}^*}$ est périodique à partir d'un certain rang (de période $n_1 - n_0$). ■

9-3 Séries absolument convergentes

9-3.1 Définition

DÉFINITION 9-3.1 *Si $(\mathbb{E}, \| \|)$ est un espace normé, et $(u_n)_{n \in \mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$, la série de terme général u_n est dite absolument convergente (en abrégé $\sum u_n$ CVA) si et seulement si la série de terme général $\|u_n\|$ converge.*

$$\sum u_n \ CVA \Leftrightarrow \sum \|u_n\| \ CV$$

Cette notion n'a vraiment d'intérêt que dans un espace complet (donc en particulier pour les séries à termes réels ou complexes), grâce au théorème :

THÉORÈME 9-3.2 Si l'espace $(\mathbb{E}, \| \|)$ est complet, toute série absolument convergente de vecteurs de $\mathbb{E}$ est convergente.

Démonstration : Qui dit espace complet dit suite de Cauchy. On montre donc que la suite des sommes partielles $S_n = \sum_{k=0}^{n} u_k$ est une

suite de Cauchy. En notant $S'_n = \sum_{k=0}^{n} \|u_k\|$, on a clairement, par inégalité triangulaire

$$\forall m > n \qquad \|S_m - S_n\| \leqslant S'_m - S'_n$$

ce qui permet de conclure aisément, puisque la suite $(S'_n)_{n\in\mathbb{N}}$ est de Cauchy puisque convergente dans $\mathbb{R}$. ■

Attention ! Il n'y a pas de réciproque. Dans $\mathbb{R}$, il existe des séries convergentes qui ne convergent pas absolument. L'exemple $u_n = \dfrac{(-1)^n}{n}$ traité en fin de section précédente le montre bien.

On peut montrer (exercice) qu'un espace normé où toute série absolument convergente est convergente est nécessairement complet.

9-3.2 Utilisation

Dans toute cette section, on suppose l'espace $(\mathbb{E}, \| \ \|)$ complet. On peut d'ailleurs, dans un premier temps, se limiter à des suites réelles ou complexes.

On comprend bien l'intérêt de la notion de convergence absolue : dans un **espace complet**, pour étudier la convergence de certaines séries, on pourra souvent se ramener à travailler avec des séries réelles positives, pour lesquelles on dispose de nombreux résultats de convergence. En particulier, si $(\mathbb{E}, \| \ \|)$ est complet, si $(u_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ est une suite de $\mathbb{E}$ et $\sum v_n$ **est une série à termes réels positifs convergente :**

$$u_n \underset{n\to+\infty}{=} O(v_n) \Rightarrow \sum u_n \text{ CVA}$$

Il en sera évidemment de même dans le cas d'une relation de négligeabilité.

La **règle de D'Alembert** peut donc en particulier être utilisée comme **règle de convergence absolue** :

$$\text{Si } \lim_{n\to+\infty} \frac{\|u_{n+1}\|}{\|u_n\|} = l \quad \text{alors} \quad \begin{cases} l < 1 \Rightarrow \sum u_n \text{ CVA} \\ l > 1 \Rightarrow \sum u_n \text{ DV grossièrement.} \end{cases}$$

le cas $l < 1$ ne pose pas de problème, le cas $l > 1$ méritant réflexion : si on affirme la divergence de $\sum u_n$, ce n'est pas parce qu'on n'a pas convergence absolue (la CVA n'est pas une condition nécessaire de convergence), mais parce qu'on sait que le terme général ne tend pas vers 0.

De même, si $(u_n)_{n\in\mathbb{N}}$ et $(u'_n)_{n\in\mathbb{N}}$ sont deux suites de vecteurs de $\mathbb{E}$ avec $u_n \underset{n\to+\infty}{\sim} u'_n$, la convergence **absolue** d'une de ces séries entrainera la convergence **absolue** de l'autre (puisque l'équivalence des termes généraux entraîne l'équivalence de leurs normes). **Mais le raisonnement ne peut être considéré comme valable que si l'expression "convergence absolue" est utilisée** (voir le contre exemple $u_n = \dfrac{(-1)^n}{n}$ et $v_n = \dfrac{(-1)^n}{n} + \dfrac{1}{n \ln n}$). C'est pourquoi il est fortement conseillé, lorsqu'on veut prouver des CVA, de travailler clairement avec les modules des termes généraux (sur $\mathbb{R}$ ou $\mathbb{C}$) ou leurs normes dans un evn complet quelconque.

REMARQUE 9-3.3 Si u_n est une suite complexe, de parties réelles et imaginaires α_n et β_n

$$u_n = \alpha_n + i\beta_n$$

les majorations $|\alpha_n| \leqslant |u_n|$, $|\beta_n| \leqslant |u_n|$ et $|u_n| \leqslant |\alpha_n| + |\beta_n|$ montrent que

$$\sum u_n \text{ CVA} \Leftrightarrow \sum \alpha_n \text{ CVA et } \sum \beta_n \text{ CVA}$$

Plus généralement, sur un espace normé de dimension finie où toutes les normes sont équivalentes, on peut travailler avec la norme "maximum des modules des coordonnées" dans une base fixe. Il est alors clair qu'une série est absolument convergente ssi les séries des coordonnées le sont toutes.

9-3.3 Majoration des restes

Il s'agit tout simplement d'un passage à la limite dans l'inégalité triangulaire. Supposons l'espace complet, et $\sum u_n$ CVA. On a alors pour $n < m$

$$\left\| \sum_{k=n+1}^{m} u_k \right\| \leqslant \sum_{k=n+1}^{m} \|u_k\|$$

En faisant tendre m vers $+\infty$, et en tenant compte de la continuité de la norme, on obtient

$$\left\| \sum_{k=n+1}^{+\infty} u_k \right\| \leqslant \sum_{k=n+1}^{+\infty} \|u_k\|$$

On en déduit que la sommation des relations de domination ou de négligeabilité se généralise de la façon suivante : si $\sum v_n$ est une **série à termes réels positifs convergente**, et si $(u_n)_{n\in\mathbb{N}} \in \mathbb{E}^{\mathbb{N}}$ (espace complet)

$$u_n \underset{n\to+\infty}{=} O(v_n) \;\Rightarrow\; \sum_{k=n+1}^{+\infty} u_k \underset{n\to+\infty}{=} O\left(\sum_{k=n+1}^{+\infty} v_k\right)$$
$$u_n \underset{n\to+\infty}{=} o(v_n) \;\Rightarrow\; \sum_{k=n+1}^{+\infty} u_k \underset{n\to+\infty}{=} O\left(\sum_{k=n+1}^{+\infty} v_k\right)$$

Attention ! La série "dominante" est à termes positifs ! On aurait des résultats analogues pour les sommes partielles, dans le cas d'une série dominante à termes positifs divergente (ici encore, il suffit d'appliquer l'inégalité triangulaire).

En réfléchissant un peu, on peut se convaincre qu'il ne faut pas espérer des résultats analogues dans le cas de l'équivalence de deux suites vectorielles (si on reprend les démonstrations du théorème **9-2**.28, c'est vraiment l'utilisation d'*encadrements* qui a amené le résultat). On pourra cependant démontrer, au coup par coup des assertions comme celle de l'exercice :

EXERCICE 9-3.4 Si v_n est une suite complexe, $u_n \geqslant 0$ le terme général d'une série convergente et s'il existe un complexe $\gamma \neq 0$ tel que

$$v_n \underset{n\to+\infty}{\sim} \gamma u_n$$

alors $\sum v_n$ CVA, et

$$\sum_{k=n+1}^{+\infty} v_k \underset{n\to+\infty}{\sim} \gamma \left(\sum_{k=n+1}^{+\infty} u_k \right)$$

9-3.4 Exemple : inverse dans une algèbre de Banach

Soit $(\mathbb{A}, +, \times, ., \| \ \|)$ une algèbre de Banach (algèbre **complète** où $\|xy\| \leqslant \|x\| \|y\|$). C'est par exemple une algèbre de matrices carrées $\mathcal{M}_p(\mathbb{K})$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, munie d'une norme subordonnée à une norme sur $\mathbb{K}^p$ (ou plus généralement l'algèbre des endomorphismes continus d'un espace de Banach). L'utilisation de séries absolument convergentes va permettre d'inverser les éléments "proches" de $1_{\mathbb{A}}$. L'idée de base est l'utilisation rigoureuse de la "formule" valable dans $\mathbb{C}$

$$|z| < 1 \Rightarrow \frac{1}{1-z} = \sum_{k=0}^{+\infty} z^k, \text{ série qui converge } \textbf{absolument}$$

THÉORÈME 9-3.5 Soit $(\mathbb{A}, +, \times, ., \| \ \|)$ une algèbre de Banach. Si $a \in \mathbb{A}$ vérifie $\|a\| < 1$, alors $1_{\mathbb{A}} - a$ est inversible dans $\mathbb{A}$ et

$$(1_{\mathbb{A}} - a)^{-1} = \sum_{k=0}^{+\infty} a^k$$

Démonstration : Comme $\left\|a^k\right\| \leqslant \|a\|^k$, l'hypothèse $\|a\| < 1$ entraîne la convergence absolue de la série $\sum a^k$. Notons

$$b = \sum_{k=0}^{+\infty} a^k$$

La démonstration sera terminée lorsqu'on aura prouvé que

$$b(1_{\mathbb{A}} - a) = (1_{\mathbb{A}} - a)b = 1_{\mathbb{A}}$$

ce qui ne peut se faire que par passage à la limite, puisque c'est ainsi qu'est construit b : pour $m \in \mathbb{N}$, on a

$$\left(\sum_{k=0}^{m} a^k \right)(1_{\mathbb{A}} - a) = (1_{\mathbb{A}} - a)\left(\sum_{k=0}^{m} a^k \right) = 1_{\mathbb{A}} - a^{m+1}$$

Comme le produit $\mathbb{A} \times \mathbb{A} \to \mathbb{A}$ est continu et que $\lim_{n\to+\infty} a^n = 0_{\mathbb{A}}$, on obtient, en faisant tendre m vers $+\infty$

$$b(1_{\mathbb{A}} - a) = (1_{\mathbb{A}} - a)b = 1_{\mathbb{A}}$$

ce qui donne bien $b = (1_{\mathbb{A}} - a)^{-1}$. ■

COROLLAIRE 9-3.6 L'ensemble $U(\mathbb{A})$ des éléments inversibles de $\mathbb{A}$ est un ouvert de $\mathbb{A}$.

Démonstration : On vient de prouver que c'est un voisinage de $1_{\mathbb{A}}$, puisqu'il contient $B(1_{\mathbb{A}}, 1[$. On en déduit facilement qu'il est voisinage de chacun de ses points : si $a \in U(\mathbb{A})$, on a

$$\forall x \in \mathbb{A} \quad a - x = a\left(1_{\mathbb{A}} - a^{-1}x\right)$$

et on est assuré de l'inversibilité de $a - x$ dès que $\|a^{-1}x\| < 1$, ce qui est certainement vérifié si $\|x\| < \dfrac{1}{\|a^{-1}\|}$. On a donc $B\left(a, \dfrac{1}{\|a^{-1}\|}\right[\subset U(\mathbb{A})$.
■

EXERCICE 9-3.7 Montrer que $a \to a^{-1}$ est un homéomorphisme de $U(\mathbb{A})$ dans lui-même. (Indication : essayer de majorer, en utilisant le développement en série, la quantité $\left\|(a-x)^{-1} - a^{-1}\right\|$ pour $\|x\|$ "assez petit").

Remarquons que, dans le cas de l'algèbre $\mathcal{M}_p(\mathbb{K})$, l'utilisation des déterminants et des comatrices nous avaient déjà permis d'obtenir le corollaire et le résultat de l'exercice précédent.

9-3.5 Espaces $l^1(\mathbb{N})$ et $l^2(\mathbb{N})$

Dans toute cette section $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On définit

$$\begin{aligned} l^1(\mathbb{N}) &= \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \mid \sum u_n \text{ CVA}\right\} \\ l^2(\mathbb{N}) &= \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \mid \sum u_n^2 \text{ CVA}\right\} \end{aligned}$$

l'espace $l^1(\mathbb{N})$ est appelé espace des suites (de $\mathbb{K}^{\mathbb{N}}$) sommables, l'espace $l^2(\mathbb{N})$ est l'espace des suites de carré sommable. Cette terminologie prendra toute sa signification dans la section consacrée aux familles sommables.

THÉORÈME 9-3.8 $l^1(\mathbb{N})$ et $l^2(\mathbb{N})$ sont deux sous-espaces vectoriels de $\mathbb{K}^{\mathbb{N}}$. Les expressions

$$\|u\|_1 = \sum_{k=0}^{+\infty} |u_k| \quad \textbf{et} \quad \|u\|_2 = \sqrt{\sum_{k=0}^{+\infty} |u_k|^2}$$

si $u = (u_n)_{n\in\mathbb{N}}$ définissent des normes respectivement sur $l^1(\mathbb{N})$ et $l^2(\mathbb{N})$.

Démonstration : La suite nulle appartient évidemment à ces deux espaces. La stabilité par homothétie est évidente. La stabilité pour la somme est claire, conséquence des inégalités

$$|u_n + v_n| \leqslant |u_n| + |v_n| \quad \text{et}$$
$$|u_n + v_n|^2 \leqslant (|u_n| + |v_n|)^2 \leqslant 2\left(|u_n|^2 + |v_n|^2\right)$$

Enfin, prouver que, pour $i = 1$ ou 2, $\| \|_i$ définit une norme sur $l^i(\mathbb{N})$ n'est pas difficile. Pour prouver l'inégalité triangulaire, il suffit de passer à la limite dans l'inégalité obtenue sur les sommes partielles des séries qui interviennent (on a en effet défini sur $\mathbb{K}^p$ deux normes qu'on a également notées $\| \|_1$ et $\| \|_2$). ■

EXERCICE 9-3.9 Montrer que $l^1(\mathbb{N}) \subset l^2(\mathbb{N})$, que l'inclusion est stricte, mais que l'injection

$$\left(l^1(\mathbb{N}), \| \|_1\right) \to \left(l^2(\mathbb{N}), \| \|_2\right) \text{ définie par } u \mapsto u$$

n'est pas continue.

THÉORÈME 9-3.10 Si $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ sont de carré sommable,

$$w = (u_n v_n)_{n\in\mathbb{N}} \in l^1(\mathbb{N})$$

et on a

$$\|w\|_1 \leqslant \|u\|_2 \|v\|_2$$

Démonstration : On a en effet

$$|u_n v_n| \leqslant \frac{1}{2}\left(|u_n|^2 + |v_n|^2\right)$$

ce qui assure l'appartenance de w à $l^1(\mathbb{N})$. De plus, l'inégalité de Schwarz dans $\mathbb{K}^m$ donne

$$\sum_{k=0}^{m} |u_k v_k| \leqslant \sqrt{\sum_{k=0}^{m} |u_k|^2}\sqrt{\sum_{k=0}^{m} |v_k|^2}$$

ce qui donne l'inégalité souhaitée, en faisant tendre m vers $+\infty$. ■

L'inégalité précédente prouve en fait la continuité de l'application bilinéaire

$$l^2(\mathbb{N}) \times l^2(\mathbb{N}) \to l^1(\mathbb{N}) \quad \text{définie par } (u, v) \mapsto w = uv$$

Lorsque $\mathbb{K} = \mathbb{R}$, la formule

$$\langle u, v\rangle = \sum_{k=0}^{+\infty} u_k v_k$$

définit un **produit scalaire** sur $l^2(\mathbb{N})$, et la norme $\| \|_2$ n'est pas autre chose que la norme associée au produit scalaire. Dans le cas $\mathbb{K} = \mathbb{C}$, pour avoir un **produit scalaire hermitien** (voir les chapitres sur les espaces préhilbertiens réels et complexes) on posera

$$\langle u, v\rangle = \sum_{k=0}^{+\infty} \overline{u_k} v_k$$

définition qui a un sens puisque cette série est absolument convergente. On a alors

$$\|u\|_2 = \sqrt{\langle u, u\rangle}$$

Nous retrouverons ces espaces (avec $\mathbb{N}$ remplacé par $\mathbb{Z}$) dans le chapitre consacré aux séries de Fourier.

EXERCICE 9-3.11 Montrer que $(l^1(\mathbb{N}), \|\ \|_1)$ et $(l^2(\mathbb{N}), \|\ \|_2)$ sont complets (Indication : on commencera par montrer que la convergence pour $\|\ \|_i$ entraîne la convergence "simple" sur $\mathbb{N}$).

EXERCICE 9-3.12 Montrer que si $l^\infty(\mathbb{N}) = \{v = (v_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \mid v \text{ est bornée}\}$, pour $v \in l^\infty(\mathbb{N})$, l'application

$$l^1(\mathbb{N}) \to \mathbb{K} \quad \text{définie par } u = (u_n)_{n\in\mathbb{N}} \mapsto \sum_{k=0}^{+\infty} v_k u_k$$

est une forme linéaire continue sur $l^1(\mathbb{N})$, de norme égale à $\|v\|_\infty$. Montrer que toute forme linéaire continue sur $l^1(\mathbb{N})$ est de cette forme. Caractériser de même les formes linéaires continues sur $l^2(\mathbb{N})$. (Réponse : $u \mapsto \sum_{k=0}^{+\infty} v_k u_k$, avec cette fois $v \in l^2(\mathbb{N})$. La norme de cette forme linéaire est ici égale à $\|v\|_2$.)

9-3.6 Exercice prolongé

Nous reprenons l'exercice **9-2.2.5**. Nous avons vu (comparaison logarithmique à une série de Riemann) que, si une série à termes positifs vérifiait

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + o\left(\frac{1}{n}\right)$$

la série convergeait pour $\alpha > 1$ et divergeait pour $\alpha < 1$ (avec divergence grossière si $\alpha < 0$). Montrer que si on a de plus

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + O\left(\frac{1}{n^2}\right)$$

alors il existe une constante $K > 0$ telle que

$$u_n \underset{n\to+\infty}{\sim} \frac{K}{n^\alpha}$$

Indication : il s'agit de montrer que la suite $n^\alpha u_n$ a une limite non nulle, c'est-à-dire que la suite $\ln(n^\alpha u_n)$ est convergente. Etudier la série télescopique associée à cette suite.

Exemple : Fonction Gamma :

Soit $a > 0$ un réel. Montrer qu'il existe un réel $\Gamma(a) > 0$ tel que

$$\frac{n!}{a(a+1)\cdots(a+n)} \underset{n\to+\infty}{\sim} \frac{\Gamma(a)}{n^a}$$

Calculer pour $p \in \mathbb{N}^*$ la valeur de $\Gamma(p)$. Montrer que $\Gamma(a+1) = a\Gamma(a)$. En admettant la formule de Stirling, calculer $\Gamma\left(\frac{1}{2}\right)$.

9-4 Séries semi-convergentes

On dit qu'une série est semi-convergente si elle converge sans converger absolument. Nous avons vu l'exemple de la série de terme général

$$u_n = \frac{(-1)^n}{n}$$

dont nous avons prouvé la convergence en **9-2.4.4** en groupant les termes deux par deux. La méthode peut être généralisée à une classe importante de séries à termes réels : les séries alternées.

9-4.1 Séries alternées

DÉFINITION 9-4.1 *Une série de terme général réel u_n est dite alternée si et seulement si $(-1)^n u_n$ a un signe constant (éventuellement à partir d'un certain rang).*

Quitte à multiplier le terme général par -1, on pourra donc supposer

$$\forall n \in \mathbb{N} \qquad u_n = (-1)^n |u_n|$$

THÉORÈME 9-4.2 Toute série alternée dont la valeur absolue du terme général décroît et tend vers 0 est convergente.

Démonstration : Bien entendu, il suffira que la suite $|u_n|$ décroisse à partir d'un certain rang pour arriver à la même conclusion. Montrons que, si S_n désigne la somme partielle d'ordre n de la série, les suites $(S_{2n})_{n\in\mathbb{N}}$ et $(S_{2n+1})_{n\in\mathbb{N}}$ forment deux suites adjacentes. On a en effet

$$\forall n \quad S_{2n+1} = S_{2n} - |u_{2n+1}| \leqslant S_{2n}$$

La suite des S_{2n} est décroissante puisque

$$S_{2n+2} - S_{2n} = |u_{2n+2}| - |u_{2n+1}| \leqslant 0$$

De même, la suite S_{2n+1} est croissante et

$$\lim_{n\to+\infty} S_{2n} - S_{2n+1} = 0$$

ce qui permet de conclure à la convergence de la série. ■

La démonstration précédente amène les précisions suivantes :

COROLLAIRE 9-4.3 Si une série converge par application du théorème précédent (dit "théorème des séries alternées"), deux sommes partielles consécutives S_n et S_{n+1} encadrent la somme de la série et donc si $S = \sum_{k=0}^{+\infty} u_k$

$$|R_n| = |S - S_n| \leqslant |u_{n+1}|$$

De plus R_n a même signe que u_{n+1} (premier terme "négligé" lorsqu'on approche S par S_n).

Démonstration : Comme S est compris entre S_n et S_{n+1}, on a clairement

$$|S - S_n| \leqslant |S_{n+1} - S_n| = |u_{n+1}|$$

De plus

$$R_n = \sum_{k=n+1}^{+\infty} u_k = \lim_{p\to+\infty} \sum_{k=n+1}^{n+2p} u_k =$$
$$\lim_{p\to+\infty} \sum_{k=1}^{p} (u_{n+2k-1} + u_{n+2k}) = \sum_{k=1}^{+\infty} (u_{n+2k-1} + u_{n+2k})$$

est somme d'une série dont le terme général est de signe constant, celui de $u_{n+1} + u_{n+2}$, c'est-à-dire celui de u_{n+1}. ■

Les résultats du corollaire précédent sont valables pour tout entier n **tel que** $(|u_k|)_{k\geqslant n}$ **soit décroissante.**

Attention ! Le théorème précédent est une condition suffisante de convergence d'une série alternée, elle n'est pas nécessaire. Par exemple, la série $\sum u_n$ CV, avec

$$u_{2p} = \frac{1}{(2p)^2} \quad \text{et} \quad u_{2p+1} = -\frac{1}{(2p+1)^3}$$

(il y a convergence absolue) mais le terme général ne décroît pas en valeur absolue. On retiendra également que, lorsque les hypothèses du théorème sont vérifiées, la convergence est assurée par groupement des termes 2 par 2, ce qui permet d'utiliser ensuite les résultats sur les séries à termes positifs.

EXEMPLE 9-4.4 Pour $\alpha > 0$, la série de terme général $u_n = \dfrac{(-1)^n}{n^\alpha}$ $(n \geqslant 1)$ est convergente. Donner un équivalent de son reste pour $n \to +\infty$.

Comme la série converge par application du théorème des séries alternées, on écrit

$$(-1)^{n+1} R_n = \sum_{k=1}^{+\infty} \left(\frac{1}{(n+2k-1)^\alpha} - \frac{1}{(n+2k)^\alpha} \right)$$

ce qui donne

$$R_{2n} = -\sum_{k=n+1}^{+\infty} \left(\frac{1}{(2k-1)^\alpha} - \frac{1}{(2k)^\alpha} \right)$$

reste d'une série dont le terme général est équivalent à $\dfrac{\alpha}{(2k)^{\alpha+1}}$. On a donc, d'après l'exemple **9-2.26**

$$R_{2n} \underset{n\to+\infty}{\sim} -\frac{\alpha}{2^{\alpha+1}} \sum_{k=n+1}^{+\infty} \frac{1}{k^{\alpha+1}} \underset{n\to+\infty}{\sim} -\frac{1}{2(2n)^\alpha}$$

On a de même

$$R_{2n+1} = R_{2n} + \frac{1}{(2n+1)^\alpha} \underset{n\to+\infty}{\sim} \frac{1}{2(2n+1)^\alpha}$$

ce qui donne finalement

$$R_n \underset{n\to+\infty}{\sim} \frac{(-1)^{n+1}}{2n^\alpha}$$

9-4.2 Méthode de décomposition

Si $u_n \underset{n\to+\infty}{\sim} v_n$ et si $\sum v_n$ CVA, on peut immédiatement conclure à la convergence absolue de $\sum u_n$. Lorsque $\sum v_n$ est seulement semi-convergente, on ne peut rien dire directement (sinon que la série de terme général u_n ne converge pas absolument). On a alors souvent recours à une décomposition de u_n, en écrivant

$$u_n = v_n + w_n$$

et il est alors clair que les séries $\sum u_n$ et $\sum w_n$ sont de même nature. Simplement, comme $w_n \underset{n\to+\infty}{=} o(v_n)$, la série de terme général w_n est peut-être plus simple à étudier, et pourrait être absolument convergente.

EXEMPLE 9-4.5 Pour $\alpha > 0$, étudier la nature de la série de terme général

$$u_n = \ln\left(1 + \frac{(-1)^n}{n^\alpha}\right) \quad \text{pour } n \geqslant 2$$

On a $u_n \underset{n\to+\infty}{\sim} \frac{(-1)^n}{n^\alpha}$ et il y a donc convergence absolue pour $\alpha > 1$. Pour $\alpha \leqslant 1$ la série de terme général $\frac{(-1)^n}{n^\alpha}$ est semi-convergente (théorème des séries alternées), on étudie u_n plus précisément, à l'aide d'un développement limité

$$u_n = \frac{(-1)^n}{n^\alpha} + w_n \text{ avec } w_n \underset{n\to+\infty}{=} -\frac{1}{2n^{2\alpha}} + o\left(\frac{1}{n^{2\alpha}}\right)$$

Ici, w_n est négative (à partir d'un certain rang), équivalente à $-\frac{1}{2n^{2\alpha}}$, et la série est donc convergente ssi $\alpha > \frac{1}{2}$.

9-4.3 Exercice : transformation d'Abel

La transformation d'Abel est une manipulation algébrique simple, qui permet de transformer la somme partielle d'une série (ou la différence de deux sommes partielles, si on veut par exemple appliquer un critère de Cauchy) en une somme partielle d'une autre série, parfois plus simple à étudier. Nous nous limiterons ici à des sommes de nombres complexes.

Si A_n et b_n sont deux suites complexes

$$\sum_{k=0}^{n} (A_{k+1} - A_k)\, b_k = -A_0 b_0 + \sum_{k=1}^{n} (b_{k-1} - b_k)\, A_k + A_{n+1} b_n$$

La vérification est simple, chaque terme A_k apparaît deux fois dans la somme de gauche (à l'exception des termes au bord A_0 et A_{n+1}). On notera l'analogie avec la formule d'intégration par parties

$$\int_{\alpha}^{\beta} g'(t) f(t)\, dt = -f(a) g(a) + \int_{\alpha}^{\beta} -f'(t) g(t)\, dt + f(b) g(b)$$

Il s'agit ici d'intégration discrète, c'est-à-dire de calcul de sommes partielles d'une série. L'opération inverse (la "dérivation") étant la correspondance qui, à la suite des sommes partielles d'une série associe le terme général, donc

$$(S_n)_{n\in\mathbb{N}} \mapsto (S_n - S_{n-1})_{n\in\mathbb{N}^*}$$

On utilisera la transformation d'Abel pour écrire autrement une somme de la forme

$$S_n = \sum_{k=0}^{n} a_k b_k$$

lorsque par exemple les sommes $A_p = \sum_{k=0}^{p} a_k$ ont des propriétés intéressantes (on peut aussi travailler avec $A_p = \sum_{k=p}^{+\infty} a_k$ si la série converge). On écrira alors

$$S_n = \sum_{k=0}^{n} (A_k - A_{k-1})\, b_k$$

en posant $A_{-1} = 0$, et on fera la transformation.

EXEMPLE 9-4.6 Discuter selon les valeurs de $\alpha > 0$ et $\theta \in [0, 2\pi]$ la nature de la série de terme général

$$u_n = \frac{e^{in\theta}}{n^\alpha} \quad (n > 0)$$

On a $|u_n| = \dfrac{1}{n^\alpha}$, et par conséquent la série est absolument convergente si $\alpha > 1$. Dans la cas $\alpha \leqslant 1$, pour $\theta = 0$ (ou 2π) la série est divergente (série de Riemann). Pour $\theta = \pi$, la série converge par application du théorème des séries alternées. Dans le cas général où $\theta \in]0, 2\pi[$, on remarque que les sommes

$$A_n = \sum_{k=0}^{n} e^{ik\theta} = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}}$$

forment une suite bornée, puisque

$$\forall n \in \mathbb{N} \qquad |A_n| \leqslant \frac{2}{|1 - e^{i\theta}|} = M$$

On aura alors

$$S_n = \sum_{k=1}^{n} \frac{e^{ik\theta}}{k^\alpha} = \sum_{k=1}^{n} (A_k - A_{k-1}) \frac{1}{k^\alpha} = -A_0 + \sum_{k=1}^{n-1} A_k \left(\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \right) + \frac{A_n}{n^\alpha}$$

Ecrite sous cette forme, on conclut aisément à la convergence de la suite S_n (donc à la convergence de $\sum u_n$). En effet $\lim_{n \to +\infty} \frac{A_n}{n^\alpha} = 0$ puisque le numérateur reste borné, et la majoration

$$\left| A_k \left(\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \right) \right| \leqslant M \left(\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \right)$$

montre la convergence absolue de la série dont la somme partielle d'ordre $n-1$ apparaît au second membre de l'égalité précédente. Il y donc convergence pour toute valeur $\alpha > 0$ lorsque $\theta \in]0, 2\pi[$.

D'une manière générale, une transformation d'Abel permettra de traiter les séries dont le terme général s'écrit $a_n b_n$ où les sommes partielles de la série de terme général a_n sont bornées, et où la suite b_n est réelle décroissante minorée (ou complexe avec $\sum (b_{n+1} - b_n)$ CVA).

REMARQUE 9-4.7 Dans l'exemple précédent, pour $\theta \in]0, 2\pi[$ et $\alpha \leqslant 1$, il n'y a pas convergence absolue de la série $\sum \frac{e^{in\theta}}{n^\alpha}$. On en déduit qu'une au moins des séries $\sum \frac{\sin(n\theta)}{n^\alpha}$ et $\sum \frac{\cos(n\theta)}{n^\alpha}$ est non absolument convergente (alors que ces deux séries sont convergentes). Les minorations

$$\frac{|\sin n\theta|}{n^\alpha} \geqslant \frac{\sin^2 n\theta}{n^\alpha} = \frac{1 - \cos 2n\theta}{2n^\alpha}$$
$$\frac{|\cos n\theta|}{n^\alpha} \geqslant \frac{\cos^2 n\theta}{n^\alpha} = \frac{1 + \cos 2n\theta}{2n^\alpha}$$

doivent permettre de démontrer que les deux séries ne convergent pas absolument (exercice).

EXERCICE 9-4.8 Nature de $\sum \frac{(-1)^n}{\sqrt[3]{n} + \cos n}$. Indication : utiliser une décomposition, puisqu'un équivalent ne permet pas de conclure. On a

$$u_n = \frac{(-1)^n}{\sqrt[3]{n}} \left(1 - \frac{\cos n}{n^{\frac{1}{3}}} + \frac{\cos^2 n}{n^{\frac{2}{3}}} - \frac{\cos^3 n}{n} + o\left(\frac{1}{n}\right) \right)$$

Pourquoi aller jusqu'à cet ordre ? Traiter ensuite chacune des séries.

9-4.4 Non "commutativité"

Les sommes de séries ne se comportent pas comme des sommes finies. Etudions par exemple la série de terme général

$$u_n = \frac{(-1)^{n-1}}{n} \text{ pour } n \geqslant 1.$$

On sait que

$$S_{2n} = \sum_{k=1}^{2n} u_k = \sum_{k=1}^{2n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\to} \ln 2 = \sum_{k=1}^{+\infty} u_k$$

Considérons à présent la suite $(v_n)_{n\in\mathbb{N}^*}$ obtenue en modifiant l'ordre des termes, en faisant apparaître un terme d'indice impair suivi de deux termes d'indices pairs de la suite u :

$$1, -\frac{1}{2}, -\frac{1}{4}, \frac{1}{3}, -\frac{1}{6}, -\frac{1}{8}, \ldots$$

soit

$$v_{3k} = \frac{1}{2k+1}, \; v_{3k+1} = -\frac{1}{4k+2} \text{ et } v_{3k+2} = -\frac{1}{4k+4}$$

La série ainsi obtenue a son terme général qui tend vers 0. Elle converge, puisqu'en regroupant les termes 3 par 3, on obtient une série convergente (voir proposition **9-1**.16) :

$$v_{3k} + v_{3k+1} + v_{3k+2} \underset{k\to+\infty}{\sim} \frac{1}{8k^2}$$

Pour calculer $\sum_{k=0}^{+\infty} v_k$, on peut grouper les termes :

$$\sum_{k=0}^{+\infty} v_k = \sum_{k=0}^{+\infty} (v_{3k} + v_{3k+1}) + v_{3k+2} = \sum_{k=0}^{+\infty} \left(\frac{1}{4k+2} - \frac{1}{4k+4} \right)$$

$$= \frac{1}{2} \sum_{k=0}^{+\infty} \left(\frac{1}{2k+1} - \frac{1}{2k+2} \right) = \frac{\ln 2}{2}$$

Le fait de modifier l'apparition des termes a divisé la somme de la série par 2 !

EXERCICE 9-4.9 Montrer que, si une série à termes réels $\sum u_n$ est semi-convergente, on peut, en modifiant l'ordre d'apparition des termes (c'est-à-dire en construisant une bijection $\varphi : \mathbb{N} \to \mathbb{N}$), obtenir

$$\sum_{k=0}^{+\infty} u_{\varphi(k)} = l$$

où l est un réel arbitrairement choisi.

Nous verrons dans la section suivante que cette situation ne peut se produire dans le cas des séries absolument convergentes, qui sont aussi "commutativement convergentes".

9-5 Notions sur les familles sommables

9-5.1 Ensemble dénombrable

DÉFINITION 9-5.1 *Un ensemble X est dit dénombrable s'il est équipotent à $\mathbb{N}$, c'est-à-dire s'il existe une bijection*

$$\varphi : \mathbb{N} \to X$$

Si on note $u_n = \varphi(n)$, cela revient aussi à dire qu'on peut "énumérer" les éléments de X, c'est-à-dire écrire

$$X = \{u_0, u_1, u_2, \ldots, u_n, \ldots\}$$

sous forme d'ensemble des valeurs d'une suite d'éléments deux à deux distincts.

Il est à remarquer que le choix d'une telle bijection semble instaurer un ordre "naturel" sur X. Il n'en est rien, puisqu'en considérant $\varphi \circ f$, avec f bijection arbitraire de $\mathbb{N}$ dans lui-même, on réalise une autre énumération des éléments de X, qui est ainsi ordonné de manière différente. Il n'y a pas de raison de privilégier l'un ou l'autre de ces ordres.

EXEMPLE 9-5.2 $\mathbb{Z}$ est dénombrable. Il suffit de définir $\varphi : \mathbb{N} \to \mathbb{Z}$ par

$$\varphi(2p) = p, \quad \varphi(2p+1) = -p-1$$

La propriété

" toute partie de $\mathbb{N}$ non vide possède un plus petit élément,"

à la base du raisonnement par récurrence, a de multiples conséquences :

1. *Toute partie d'un ensemble dénombrable est finie ou dénombrable :* transportant le problème par bijection, il suffit de le démontrer pour $\mathbb{N}$. Ne souhaitant pas faire de théorie formelle des ensembles, nous prendrons ici le terme "fini" au sens intuitif (ce qui est loin d'être satisfaisant, il faut le reconnaître). Si $\emptyset \neq A \subset \mathbb{N}$, on définit

   $$a_0 = \min A$$

   Si A ne contient qu'un élément, on s'arrête. Sinon, on pose

   $$a_1 = \min(A - \{a_0\})$$

   puis si c'est possible

   $$a_p = \min(A - \{a_0, a_1, \ldots, a_{p-1}\})$$

   Si A est fini, le processus s'arrête, sinon on a clairement (?)

   $$A = \{a_0, a_1, \ldots, a_n, \ldots\}$$

2. *Si A est dénombrable et $f : A \to X$ est une application à valeurs dans un ensemble quelconque, alors $f(A)$ est fini ou dénombrable.* En effet, si $A = \mathbb{N}$, (on peut évidemment se ramener à cette situation) l'application

$$f(A) \to \mathbb{N} \quad x \mapsto \min\{i \in \mathbb{N} \mid f(i) = x\}$$

réalise une injection de $f(A)$ dans $\mathbb{N}$, donc une bijection entre $f(A)$ et une partie de $\mathbb{N}$.

3. *$\mathbb{N} \times \mathbb{N}$ et $\mathbb{Q}$ sont dénombrables.* L'application $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ définie par $(p,q) \mapsto 2^p 3^q$ est en effet injective. On peut d'ailleurs réaliser une bijection de $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ en considérant $\mathbb{N} \times \mathbb{N}$ comme un tableau à double entrée, et en numérotant les éléments (p,q) de ce tableau selon l'ordre lexicographique sur $(p+q,p)$, c'est-à-dire en ordonnant

$$(0,0),(0,1),(1,0),(0,2),(1,1),(2,0),(0,3),(1,2) \text{ etc...}$$

et en posant donc

$$\varphi(p,q) = p + \frac{(p+q)(p+q+1)}{2}$$

De même, l'application $\mathbb{Q} \to \mathbb{N}$ définie sur des rationnels écrits sous forme irréductibles par

$$\forall (p,q) \in \mathbb{N}^* \times \mathbb{N}^* \quad \varphi(0) = 1, \quad \varphi\left(\frac{p}{q}\right) = 2^p 3^q \quad \text{et} \quad \varphi\left(-\frac{p}{q}\right) = 5^p 7^q$$

est injective, et prouve que $\mathbb{Q}$ est dénombrable. Ce résultat semble étrange au premier abord, parce qu'on pense immédiatement à la structure d'ordre sur $\mathbb{Q}$. La structure d'ordre usuelle sur $\mathbb{Q}$ est dite "dense", parce qu'entre deux rationnels distincts il y a une infinité de rationnels. Il faut bien comprendre qu'une bijection entre $\mathbb{N}$ et $\mathbb{Q}$ induit un ordre discret sur $\mathbb{Q}$, qui n'a rien à voir avec cet ordre usuel.

4. *Un produit fini d'ensembles dénombrables est dénombrable.* Il suffit de le montrer pour $\mathbb{N}^p$, ce qui se fait par récurrence sur p à partir du résultat 3).

5. *Une réunion finie ou dénombrable d'ensembles dénombrables est dénombrable.* Montrons le sur une suite $(A_n)_{n\in\mathbb{N}}$ d'ensembles dénombrables, dont les éléments sont énumérés

$$A_i = \{a_j^i,\ j \in \mathbb{N}\}$$

L'application

$$\mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n \quad \text{définie par} \quad (i,j) \mapsto a_j^i$$

est évidemment surjective. Comme les A_i sont infinis, le résultat 3) permet de conclure. On démontrerait de même qu'*une union dénombrable d'ensembles finis est finie ou dénombrable.*

6. $\mathbb{R}$ *n'est pas dénombrable.* Ici, l'intuition est mise à rude épreuve : la première idée qu'on puisse avoir de l'infini c'est sans doute sous la forme d'une suite d'objets deux à deux distincts (vision déjà idéalisée, car non intuitive). $\mathbb{R}$ est un exemple d'ensemble ayant "beaucoup plus" d'éléments que $\mathbb{N}$, résultat pas si évident que cela (si on pense que $\mathbb{Q}$ est équipotent à $\mathbb{N}$). Il suffit de montrer que $[0,1]$ est non dénombrable. Si c'était le cas, on aurait

$$[0,1] = \{x_n,\ n \in \mathbb{N}\}$$

Il existe au moins un des trois segments $\left[0,\frac{1}{3}\right]$, $\left[\frac{1}{3},\frac{2}{3}\right]$ et $\left[\frac{2}{3},1\right]$ qui ne contient pas x_0. Appelons K_0 ce segment. En le coupant en trois, on obtient un nouveau segment K_1 qui ne contient ni x_0 ni x_1. En réitérant le procédé, on forme une suite $(K_n)_{n\in\mathbb{N}}$ de segments emboîtés, de longueurs $\frac{1}{3^{n+1}}$ et d'intersection vide, ce qui est contraire à la compacité de $[0,1]$.

7. *L'ensemble des parties de* $\mathbb{N}$ *est non dénombrable.* C'est un cas particulier d'un résultat dû à Cantor, qui dit que, pour X ensemble quelconque,

$$X \text{ et } \mathcal{P}(X) \text{ ne sont pas équipotents}$$

(résultat "évident" si X est fini, puisque $2^p > p$ pour tout entier naturel p). Il suffit pour cela d'envisager une application $f : X \to \mathcal{P}(X)$ et de considérer $A = \{x \in X \mid x \notin f(x)\}$ pour voir que f ne peut être surjective.

EXERCICE 9-5.3 En utilisant le développement propre des réels de $]0,1[$ en base 2 (c'est-à-dire le développement qui ne se termine pas par une suite de digits tous égaux à 1), montrer que $]0,1[$ est équipotent à l'ensemble des parties infinies de $\mathbb{N}$.

EXERCICE 9-5.4 Tout ensemble infini contient un sous-ensemble dénombrable (c'est "évident" ?). Montrer que, si A est un ensemble infini non dénombrable et $D \subset A$ en est une partie dénombrable, alors A et $A-D$ sont équipotents. (Indication : considérer une partie dénombrable D' de $A-D$ et réaliser une bijection de $D\cup D'$ sur D'). Déduire alors de l'exercice précédent que $]0,1[$ est équipotent à $\mathcal{P}(\mathbb{N})$.

9-5.2 Famille sommable de réels positifs

Dans tout ce qui suit, I est un ensemble dénombrable d'indices. Il importe de bien comprendre qu'il n'y a alors pas d'ordre privilégié sur I (bien que le choix d'une bijection particulière $\mathbb{N} \to I$ permette d'ordonner I). On sait qu'il existe alors des suites de parties finies $J_n \subset I$, croissantes pour l'inclusion ($J_n \subset J_{n+1}$) telles que $I = \bigcup_{n\in\mathbb{N}} J_n$. On notera alors

$$I = \bigcup_{\substack{\nearrow \\ n\in\mathbb{N}}} J_n$$

pour résumer cette situation. On peut (par exemple) prendre une bijection $\varphi : \mathbb{N} \to I$ et poser $J_n = \varphi([0,a_n])$, où a_n est une suite croissante non stationnaire d'entiers.

Notation : Si $(\mathbb{E}, \| \ \|)$ est un espace normé et si $u = (u_i)_{i \in I}$ est une famille de vecteurs de $\mathbb{E}$ indexée par I, **pour** $J \subset I$ **partie finie de** I, on notera

$$S_J(u) = \sum_{j \in J} u_j$$

Nous nous limiterons pour l'essentiel à $\mathbf{E} = \mathbb{R}$ ou $\mathbb{C}$, et dans un premier temps nous travaillerons avec des familles de réels positifs.

DÉFINITION 9-5.5 *Soit $u = (u_i)_{i \in I} \in (\mathbb{R}^+)^I$ une famille de réels positifs indexée par un ensemble dénombrable I. On dit que cette famille est sommable si et seulement si*

$$\left\{ S_J(u) = \sum_{j \in J} u_j \ , \ J \text{ partie finie de } I \right\}$$

est majoré (dans $\mathbb{R}^+$). Lorsque c'est le cas, on définit la somme de cette famille par

$$S(u) = sup\,\{S_J(u) \ , \ J \text{ partie finie de } I\}$$

On notera aussi

$$S(u) = \sum_{i \in I} u_i$$

Cette somme est définie comme borne supérieure d'un sous ensemble de $\mathbb{R}^+$. On pourra bien sûr approcher cette borne supérieure, et plus précisément la déterminer comme limite d'une suite croissante de réels positifs :

THÉORÈME 9-5.6 S'il existe une suite croissante (J_n) de parties finies de I dont la réunion est égale à I telle que $(S_{J_n}(u))_{n \in \mathbb{N}}$ soit une suite majorée, alors la famille u est sommable et

$$S(u) = \sup_n S_{J_n}(u) = \lim_{n \to +\infty} S_{J_n}(u)$$

Réciproquement, si u est sommable, cette égalité est valable pour toute suite (J_n) de parties finies de I avec $I = \bigcup_{\nearrow\, n \in \mathbb{N}} J_n$.

Démonstration : Si $J \subset I$ est une partie finie, on a $J \subset J_n$ pour n suffisamment grand. Comme les réels manipulés sont positifs

$$\exists N \in \mathbb{N} \quad \forall n \geqslant N \quad S_J(u) \leqslant S_{J_n}(u) \leqslant \lim_{n \to +\infty} S_{J_n}(u)$$

Comme cette inégalité est valable pour tout J, on a la sommabilité de u avec

$$S(u) \leqslant \lim_{n \to +\infty} S_{J_n}(u)$$

avec égalité, puisque l'inégalité inverse est évidente (définition de $S(u)$). Réciproquement, si u est sommable, le raisonnement précédent peut être fait pour toute suite (J_n) avec $I = \bigcup_{\nearrow\, n \in \mathbb{N}} J_n$. ■

Le cas $I = \mathbb{N}$ est important, et on fait le lien avec la notion de série :

THÉORÈME 9-5.7 Une suite $u = (u_n)_{n\in\mathbb{N}} \in (\mathbb{R}^+)^{\mathbb{N}}$ de réels positifs est sommable si et seulement si la série de terme général u_n converge et on a alors

$$S(u) = \sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$$

Démonstration : Il suffit d'appliquer le théorème précédent avec $J_n = [0, n]$. ∎

Dans l'égalité précédente, deux symboles différents interviennent : $\sum_{n=0}^{+\infty} u_n$ somme de la **série** de terme général u_n qui, dans sa définition, semble tenir compte de l'ordre d'apparition des termes et $\sum_{n\in\mathbb{N}} u_n$ somme de la **famille** (sommable) $(u_n)_{n\in\mathbb{N}}$, qui dans sa définition ne tient pas compte de l'ordre "naturel" sur $\mathbb{N}$. Dans le cas de réels positifs, une modification de l'ordre d'apparition des termes de la série ne change pas la nature ou la somme de la série. Nous y reviendrons ultérieurement dans le cadre des séries absolument convergentes. Il faut noter la différence avec la situation rencontrée à la section **9-4.4**.

9-5.3 Famille sommable de complexes

On considère une famille $u = (u_i)_{i\in I} \in \mathbb{C}^I$.

DÉFINITION 9-5.8 *La famille $u = (u_i)_{i\in I} \in \mathbb{C}^I$ est dite sommable ssi la famille des modules $(|u_i|)_{i\in I}$ l'est.*

Pour vérifier la sommabilité d'une famille de complexes, on se ramènera donc toujours à travailler avec des familles de réels positifs.

THÉORÈME 9-5.9 (ET DEFINITION) Soit $u = (u_i)_{i\in I} \in \mathbb{C}^I$ une famille complexe sommable. Pour toute suite (J_n) de sous-familles finies de I avec $I = \bigcup\limits_{\substack{\nearrow \\ n\in\mathbb{N}}} J_n$, la suite $(S_{J_n}(u))_{n\in\mathbb{N}}$ converge. Sa limite ne dépend pas de la suite (J_n) choisie, et on posera

$$S(u) = \sum_{i\in I} u_i = \lim_{n\to+\infty} \sum_{i\in J_n} u_i$$

Démonstration : On montre que la suite $(S_{J_n}(u))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{C}$. Par inégalité triangulaire, on a évidemment

$$n < m \Rightarrow |S_{J_m}(u) - S_{J_n}(u)| \leqslant S_{J_m}(|u|) - S_{J_n}(|u|)$$

ce qui permet de conclure, puisqu'on sait par définition que la suite $(S_{J_n}(|u|))_{n\in\mathbb{N}}$ est convergente (famille sommable de réels positifs).

Pour voir que la limite ne dépend pas de la suite (J_n) choisie, considérons deux telles suites (J_n) et (K_n) et posons $L_n = J_n \cap K_n$. Cette suite est non vide à partir d'un certain rang, et vérifie

$$I = \bigcup_{\substack{\nearrow \\ n\in\mathbb{N}}} L_n$$

(en effet, si $i \in I$, il existe des indices n_1 et n_2 avec $i \in J_{n_1}$ et $i \in K_{n_2}$, donc $i \in L_{\max(n_1,n_2)}$ puisque les suites sont croissantes pour l'inclusion). On a toujours, par inégalité triangulaire

$$|S_{J_n}(u) - S_{L_n}(u)| \leqslant S_{J_n}(|u|) - S_{L_n}(|u|)$$

qui est une suite tendant vers 0 puisque

$$\lim_{n\to+\infty} S_{J_n}(|u|) = \lim_{n\to+\infty} S_{L_n}(|u|) = S(|u|)$$

On en déduit que

$$\lim_{n\to\infty} S_{J_n}(u) = \lim_{n\to\infty} S_{J_n\cap K_n}(u) = \lim_{n\to\infty} S_{K_n}(u)$$

pour des raisons de symétrie. ∎

REMARQUE 9-5.10 On aurait pu définir autrement cette somme, en regardant d'abord le cas où $(u_i)_{i\in I}$ est une **famille réelle**. On pose alors

$$u_i^+ = \max(u_i, 0) \text{ et } u_i^- = \max(-u_i, 0)$$

("parties positives et négatives de la famille $(u_i)_{i\in I}$"), ce qui donne

$$u_i = u_i^+ - u_i^- \quad \text{et} \quad |u_i| = u_i^+ + u_i^-$$

Les encadrements

$$0 \leqslant u_i^+ \leqslant |u_i| \quad \text{et} \quad 0 \leqslant u_i^- \leqslant |u_i|$$

montrent que les familles (u_i^+) et (u_i^-) sont sommables, ce qui prouve l'existence de

$$\lim_{n\to+\infty} S_{J_n}(u) = \lim_{n\to+\infty} S_{J_n}(u^+) - \lim_{n\to+\infty} S_{J_n}(u^-)$$

soit finalement $S(u) = S(u^+) - S(u^-)$. Lorsque u est complexe, on pourra séparer les parties réelles et imaginaires, qui formeront deux familles sommables, puisque

$$|\mathrm{Re}\, u_i| \leqslant |u_i| \quad \text{et} \quad |\mathrm{Im}\, u_i| \leqslant |u_i|$$

et $S(u)$ vaudra simplement

$$S(u) = \left(S\left((\mathrm{Re}\, u)^+\right) - S\left((\mathrm{Re}\, u)^-\right)\right) + i\left(S\left((\mathrm{Im}\, u)^+\right) - S\left((\mathrm{Im}\, u^-)\right)\right)$$

Il est plus intéressant d'utiliser le critère de Cauchy, puisque le résultat se généralisera à une **famille d'un espace normé complet** : une famille $(u_i) \in \mathbb{E}^I$ sera dite (absolument) sommable ssi la famille des normes $(\|u_i\|)_{i\in I}$ l'est. Le résultat du théorème précédent persiste alors, avec une démonstration analogue.

Dans le cas particulier $I = \mathbb{N}$, le théorème **9-5**.7 donne

PROPOSITION 9-5.11 Une suite à termes complexes (ou à valeurs dans un espace de Banach) $u = (u_n)_{n\in\mathbb{N}}$ est sommable ssi la série de terme général u_n est absolument convergente, et on a alors

$$S(u) = \sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$$

Ici encore, l'ordre d'apparition des termes n'intervient pas dans la définition de $S(u)$. On en déduit

COROLLAIRE 9-5.12 Toute série complexe (ou dans un Banach) absolument convergente est commutativement convergente, c'est-à-dire que

$$\sum u_n \text{ CVA} \Rightarrow \forall \varphi \textbf{ bijective } \mathbb{N} \to \mathbb{N} \quad \sum u_{\varphi(n)} \text{ CVA et} \sum_{k=0}^{+\infty} u_{\varphi(k)} = \sum_{k=0}^{+\infty} u_k$$

Démonstration : La famille $(u_n)_{n\in\mathbb{N}}$ est sommable. La convergence absolue de $\sum u_{\varphi(n)}$ en découle puisque

$$\forall N \in \mathbb{N} \quad \sum_{k=0}^{N} \left|u_{\varphi(k)}\right| \leqslant S(|u|)$$

De plus, comme φ est bijective, on a $\mathbb{N} = \bigcup_{\nearrow} \varphi([0, n])$, et par conséquent

$$\lim_{n\to+\infty} \sum_{k=0}^{n} u_{\varphi(k)} = S(u) \quad \blacksquare$$

EXERCICE 9-5.13 Montrer que $(u_n)_{n\in\mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}}$ est sommable si et seulement si les séries

$$\sum_{n=0}^{+\infty} u_n \quad \text{et} \quad \sum_{n=1}^{+\infty} u_{-n}$$

sont absolument convergentes, et montrer qu'alors

$$S(u) = \sum_{n\in\mathbb{Z}} u_n = \sum_{n=0}^{+\infty} u_n + \sum_{n=1}^{+\infty} u_{-n} = \lim_{n\to+\infty} \sum_{k=-n}^{n} u_k$$

Le contre-exemple $u_k = \dfrac{k}{k^2+1}$ (suite impaire) montre que l'existence de la dernière limite n'implique pas la sommabilité de $(u_n)_{n\in\mathbb{Z}}$.

9-5.4 Suites doubles sommables

On se limite dans cette section au cas particulier $I = \mathbb{N} \times \mathbb{N}$. Conformément aux définitions précédentes, nous étudierons d'abord le cas de familles de réels positifs.

9-5.4.1 Cas de réels positifs

On considère une suite double $u = (u_{p,q})_{(p,q)\in\mathbb{N}\times\mathbb{N}} \in (\mathbb{R}^+)^{\mathbb{N}\times\mathbb{N}}$.

THÉORÈME 9-5.14 La famille $u = (u_{p,q})_{(p,q)\in\mathbb{N}\times\mathbb{N}} \in (\mathbb{R}^+)^{\mathbb{N}\times\mathbb{N}}$ est sommable si et seulement si

$$\begin{cases} \forall q \in \mathbb{N} \quad \textbf{la série de terme général } (u_{p,q})_{p\in\mathbb{N}} \textbf{ est convergente} \\ \textbf{La série de terme général } v_q = \sum_{p=0}^{+\infty} u_{p,q} \textbf{ converge} \end{cases}$$

On a alors

$$\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}} u_{p,q} = \sum_{q=0}^{+\infty} v_q = \sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty} u_{p,q}\right)$$

Comme les indices p et q jouent de rôles symétriques, on pourra intervertir l'ordre des sommations et on aura aussi

$$\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}} u_{p,q} = \sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty} u_{p,q}\right)$$

Démonstration : Supposons la suite double sommable. Notons

$$S = S(u) = \sum_{(p,q)\in\mathbb{N}\times\mathbb{N}} u_{p,q}$$

Comme cette somme est borne supérieure des sommes sur les parties finies de $\mathbb{N} \times \mathbb{N}$, on a , pour m et n entiers

$$\sum_{q=0}^{n}\left(\sum_{p=0}^{m} u_{p,q}\right) \leqslant S \qquad (*)$$

Comme tous les termes sont positifs, on a en particulier

$$\forall q \quad \forall m \qquad \sum_{p=0}^{m} u_{p,q} \leqslant S$$

A q fixé, les sommes partielles de la série (à termes positifs) de terme général $(u_{p,q})_{p\in\mathbb{N}}$ sont majorées par S. Il y a donc convergence de cette série. En posant

$$v_q = \sum_{p=0}^{+\infty} u_{p,q}$$

on a, en faisant tendre m vers l'infini dans $(*)$

$$\forall n \in \mathbb{N} \quad \sum_{q=0}^{n} v_q \leqslant S$$

majorations qui prouvent la convergence de la série de terme général (positif) v_q, avec

$$\sum_{q=0}^{+\infty} v_q \leqslant S$$

L'inégalité inverse est évidente, puisque, si J est une partie finie de $\mathbb{N} \times \mathbb{N}$, on peut trouver des entiers m et n avec $J \subset [0, m] \times [0, n]$ et alors

$$\sum_{(p,q)\in J} u_{p,q} \leqslant \sum_{q=0}^{n} \left(\sum_{p=0}^{m} u_{p,q}\right) \leqslant \sum_{q=0}^{n} \left(\sum_{p=0}^{+\infty} u_{p,q}\right) = \sum_{q=0}^{n} v_q \leqslant \sum_{q=0}^{+\infty} v_q$$

ce qui donne bien, par définition de S

$$S \leqslant \sum_{q=0}^{+\infty} v_q$$

Lorsqu'on suppose que, réciproquement, les séries définissant v_q et $\sum_{q=0}^{+\infty} v_q$ convergent, la fin du raisonnement précédent peut être recopiée et prouve la sommabilité de la suite double. ■

EXEMPLE 9-5.15 Discuter selon les valeurs du paramètre $\alpha > 0$ la sommabilité de

$$\left(\frac{1}{(p+q)^{\alpha}}\right)_{p,q\in\mathbb{N}^*}$$

La condition $\alpha > 1$ est nécessaire pour avoir convergence de la série à q fixé. Si cette condition est vérifiée

$$\int_{q+1}^{+\infty} \frac{dt}{t^{\alpha}} \leqslant \sum_{p=1}^{+\infty} \frac{1}{(p+q)^{\alpha}} \leqslant \int_{q}^{+\infty} \frac{dt}{t^{\alpha}}$$

encadrement qui permet de conclure à la sommabilité si et seulement si $\alpha > 2$.

EXERCICE 9-5.16 Pour $x > 1$, on définit

$$\zeta(x) = \sum_{n=1}^{+\infty} \frac{1}{n^x}$$

Convergence et calcul de

$$\sum_{n=2}^{+\infty} (\zeta(n) - 1)$$

9-5.4.2 Suites doubles de complexes

Nous considèrerons des suites doubles $u = (u_{p,q})_{(p,q)\in\mathbb{N}\times\mathbb{N}} \in (\mathbb{C})^{\mathbb{N}\times\mathbb{N}}$, mais l'étude pourrait être menée dans un espace de Banach quelconque, en remplaçant le module par la norme. Comme la sommabilité fait intervenir la suite $(|u_{p,q}|)_{(p,q)\in\mathbb{N}\times\mathbb{N}}$, le théorème **9-5**.14 donne

THÉORÈME 9-5.17 Si $u = (u_{p,q})_{(p,q)\in\mathbb{N}\times\mathbb{N}}$ **est une suite double complexe, cette suite est sommable si et seulement si**

$$\sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}|u_{p,q}|\right) \text{ existe.}$$

De manière équivalente, il suffit aussi de vérifier que

$$\sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty}|u_{p,q}|\right) < +\infty$$

Lorsqu'une de ces conditions est vérifiée, on pourra "intervertir l'ordre des sommations". Plus précisément :

THÉORÈME 9-5.18 Si $(u_{p,q})_{(p,q)\in\mathbb{N}\times\mathbb{N}}$ **est sommable, les séries intervenant dans les définitions de**

$$\sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}u_{p,q}\right) \text{ et } \sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty}u_{p,q}\right)$$

sont toutes (absolument) convergentes, et

$$\sum_{(p,q)\in\mathbb{N}^2}u_{p,q} = \sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}u_{p,q}\right) = \sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty}u_{p,q}\right)$$

Démonstration : L'hypothèse de sommabilité entraîne évidemment, à q fixé, la convergence absolue de la série $\sum u_{p,q}$. Comme de plus

$$\left|\sum_{p=0}^{+\infty}u_{p,q}\right| \leqslant \sum_{p=0}^{+\infty}|u_{p,q}|$$

la sommabilité de la famille $(|u_{p,q}|)$ et le théorème **9-5**.14 entraîne la convergence absolue de la série $\sum w_q$, avec $w_p = \sum_{p=0}^{+\infty} u_{p,q}$. Il reste à prouver que

$$\sum_{(p,q)\in\mathbb{N}^2}u_{p,q} = \sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}u_{p,q}\right)$$

(par changement du rôle joué par les indices p et q, on aura en même temps la possibilité d'intervertir les sommations). Si on pose

$S = \sum_{(p,q)\in\mathbb{N}^2} u_{p,q}$, cela revient à montrer, revenant à la définition de la convergence d'une série, que

$$\lim_{n\to+\infty}\left|S-\sum_{q=0}^{n}\left(\sum_{p=0}^{+\infty}u_{p,q}\right)\right|=0$$

On a, par inégalité triangulaire

$$\left|S-\sum_{q=0}^{n}\left(\sum_{p=0}^{+\infty}u_{p,q}\right)\right|\leqslant\left|S-\sum_{q=0}^{n}\left(\sum_{p=0}^{n}u_{p,q}\right)\right|+\left|\sum_{q=0}^{n}\left(\sum_{p=n+1}^{+\infty}u_{p,q}\right)\right|$$

Montrons que les deux termes intervenant dans le terme de droite tendent vers 0. C'est évident pour le premier, puisque $[0,n]\times[0,n]$ forme une suite croissante de parties dont la réunion est $\mathbb{N}^2$. Pour le second, on a

$$\left|\sum_{q=0}^{n}\left(\sum_{p=n+1}^{+\infty}u_{p,q}\right)\right|\leqslant\sum_{q=0}^{n}\left|\left(\sum_{p=n+1}^{+\infty}u_{p,q}\right)\right|$$
$$\leqslant\sum_{q=0}^{n}\left(\sum_{p=n+1}^{+\infty}|u_{p,q}|\right)=\sum_{p=n+1}^{+\infty}\left(\sum_{q=0}^{n}|u_{p,q}|\right)$$

(inégalité triangulaire d'abord, puis somme de $n+1$ séries convergentes). On conclut ensuite, en majorant

$$\sum_{p=n+1}^{+\infty}\left(\sum_{q=0}^{n}|u_{p,q}|\right)\leqslant\sum_{p=n+1}^{+\infty}\left(\sum_{q=0}^{+\infty}|u_{p,q}|\right)$$

reste d'une série convergente (théorème **9-5**.14) et tend vers 0 pour $n\to+\infty$. ■

EXERCICE 9-5.19 Pour $|z|<1$, avec $z\in\mathbb{C}$, établir

$$\sum_{n=1}^{+\infty}\frac{z^n}{1-z^n}=\sum_{n=1}^{+\infty}d(n)\,z^n$$

où $d(n)$ désigne le nombre de diviseurs positifs de n. (Il s'agit de prouver la convergence des deux séries et l'égalité de leurs sommes).

REMARQUE 9-5.20 Si $(u_{p,q})_{(p,q)\in\mathbb{N}^2}$ est une suite double **quelconque** de complexes, on peut avoir existence et inégalité des deux termes

$$\sum_{q=0}^{+\infty}\left(\sum_{p=0}^{+\infty}u_{p,q}\right)\neq\sum_{p=0}^{+\infty}\left(\sum_{q=0}^{+\infty}u_{p,q}\right)$$

On peut voir par exemple que, pour $(u_{p,q})_{(p,q)\in\mathbb{N}^*\times\mathbb{N}^*}$ définie par

$$u_{p,q}=\frac{1}{p^2-q^2}\text{ pour }p\neq q\text{ et }u_{p,q}=0\text{ si }p=q$$

on a

$$\sum_{p=1}^{+\infty} u_{p,q} = \lim_{N\to+\infty} \sum_{\substack{p=1 \\ p\neq q}}^{N} \left(\frac{1}{p-q} - \frac{1}{p+q}\right) \frac{1}{2q}$$

En notant S_k la somme partielle d'ordre k de la série harmonique, on a

$$\sum_{p=1}^{+\infty} u_{p,q} = \frac{1}{2q} \lim_{N\to+\infty} \left(\frac{1}{q} - S_{N+q} + S_{N-q}\right) = \frac{1}{2q^2}$$

en encadrant $S_{N+q} - S_{N-q}$ par deux intégrales. On obtient

$$\sum_{q=1}^{+\infty} \left(\sum_{p=1}^{+\infty} u_{p,q}\right) = \frac{1}{2} \sum_{q=1}^{+\infty} \frac{1}{q^2}$$

En permutant les rôles joués par p et q, le résultat est évidemment opposé !

9-5.4.3 Produit de deux suites sommables. Produit de Cauchy de deux séries absolument convergentes

THÉORÈME 9-5.21 (ET DEFINITION) **Si $a = (a_p)_{p\in\mathbb{N}}$ et $b = (b_q)_{q\in\mathbb{N}}$ sont deux suites complexes sommables, la suite double $(a_p b_q)_{(p,q)\in\mathbb{N}^2}$ est sommable. On l'appelle famille produit des familles a et b et on a**

$$\sum_{(p,q)\in\mathbb{N}^2} a_p b_q = \left(\sum_{p\in\mathbb{N}} a_p\right) \left(\sum_{q\in\mathbb{N}} b_q\right)$$

Démonstration : La suite double positive $(|a_p b_q|)_{(p,q)\in\mathbb{N}^2}$ est sommable, puisque

$$\sum_{q=0}^{+\infty} \left(\sum_{p=0}^{+\infty} |b_q|\,|a_p|\right) \text{ converge.}$$

Le résultat en découle, avec

$$\sum_{(p,q)\in\mathbb{N}^2} a_p b_q = \sum_{p=0}^{+\infty} \left(\sum_{q=0}^{+\infty} a_p b_q\right) = \left(\sum_{p=0}^{+\infty} a_p\right) \left(\sum_{q=0}^{+\infty} b_q\right) = \left(\sum_{p\in\mathbb{N}} a_p\right) \left(\sum_{q\in\mathbb{N}} b_q\right) \blacksquare$$

REMARQUE 9-5.22 On aurait un résultat analogue en définissant le produit de deux familles complexes sommables $(a_i)_{i\in I}$ et $(b_j)_{j\in J}$ comme étant $(a_i b_j)_{(i,j)\in I\times J}$ avec

$$\sum_{(i,j)\in I\times J} a_i b_j = \left(\sum_{i\in I} a_i\right) \left(\sum_{j\in J} b_j\right)$$

Pour faire la démonstration, il suffirait de prendre des énumérations de I et J et de se ramener à la situation précédente.

REMARQUE 9-5.23 En travaillant avec des espaces normés complets, et en remplaçant le produit par une application bilinéaire continue, on pourrait généraliser le théorème précédent :

Si $\mathbb{E}, \mathbb{F}, \mathbb{G}$ sont trois espaces de Banach, si $B : \mathbb{E} \times \mathbb{F} \to \mathbb{G}$ est bilinéaire continue et si $(u_p) \in \mathbb{E}^{\mathbb{N}}$ et $(v_q) \in \mathbb{F}^{\mathbb{N}}$ sont deux suites (absolument) sommables, la famille

$$(B(u_p, v_q))_{(p,q)\in\mathbb{N}^2} \in \mathbb{G}^{\mathbb{N}^2}$$

est sommable, et

$$B\left(\sum_{p\in\mathbb{N}} u_p, \sum_{q\in\mathbb{N}} v_q\right) = \sum_{(p,q)\in\mathbb{N}^2} B(u_p, v_q)$$

La démonstration n'est pas difficile. On utilise l'existence d'une constante K telle que

$$\forall (x,y) \in \mathbb{E} \times \mathbb{F} \quad \|B(x,y)\| \leqslant K \|x\| \|y\|$$

ce qui traduit la continuité de B, et le fait que $J_n = [0,n] \times [0,n]$ forme une suite croissante d'union $\mathbb{N}^2$, ce qui permet de passer à la limite dans l'égalité

$$B\left(\sum_{p=0}^{n} u_p, \sum_{q=0}^{n} v_q\right) = \sum_{(p,q)\in J_n} B(u_p, v_q)$$

En ce qui nous concerne, nous utiliserons cette notion de produit presque exclusivement dans le cadre des séries absolument convergentes, avec la notion de produit de Cauchy.

THÉORÈME 9-5.24 (ET DEFINITION) Si $\sum a_n$ et $\sum b_n$ sont deux séries à termes complexes absolument convergentes, on appelle série produit de Cauchy de ces deux séries la série de terme général

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{\substack{(p,q)\in\mathbb{N}^2 \\ p+q=n}} a_p b_q$$

Cette série est absolument convergente, de somme

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right)$$

Démonstration : La famille $(a_p b_q)$ est sommable, d'après ce qui précède. Si on pose

$$J_n = \left\{(p,q) \in \mathbb{N}^2 \mid 0 \leqslant p+q \leqslant n\right\}$$

on construit une suite croissante de parties finies de $\mathbb{N}^2$ de réunion égale à $\mathbb{N}^2$. On a donc

$$\lim_{n\to+\infty} \sum_{(p,q)\in J_n} a_p b_q = \lim_{n\to+\infty} \sum_{k=0}^{n} c_k = \sum_{(p,q)\in\mathbb{N}^2} a_p b_q = \left(\sum_{p\in\mathbb{N}} a_p\right)\left(\sum_{q\in\mathbb{N}} b_q\right)$$

ce qui prouve la convergence de la série et détermine sa somme. La convergence absolue est simplement conséquence de la majoration

$$|c_n| \leqslant \sum_{k=0}^{n} |a_k| \, |b_{n-k}|$$

la série majorante étant convergente (produit de Cauchy de deux séries convergentes à termes positifs). ■

Nous donnerons au début de la section suivante l'exemple fondamental de la fonction exponentielle. Terminons par deux remarques :

REMARQUE 9-5.25 Si $\sum a_n$ et $\sum b_n$ sont deux séries à termes complexes semi-convergentes, la convergence de $\sum c_n$ avec

$$c_n = \sum_{\substack{(p,q) \\ p+q=n}} a_p b_q$$

n'est plus assurée. Par exemple, pour $a_n = b_n = \dfrac{(-1)^n}{\sqrt{n}}$ $(n \geqslant 1)$ séries alternées convergentes,

$$c_n = (-1)^n \sum_{k=1}^{n-1} \frac{1}{\sqrt{k\,(n-k)}}$$

ne tend pas vers 0 (exercice) et donc $\sum c_n$ DV. On peut cependant prouver (exercice) que, si l'une des séries est ACV et l'autre converge, la série produit de Cauchy converge et a pour somme le produit des sommes (théorème de Mertens).

REMARQUE 9-5.26 Le théorème précédent se généralise dans le cas de séries absolument convergentes dans des espaces de Banach et d'une application bilinéaire continue : la série de terme général

$$c_n = \sum_{k=0}^{n} B\,(a_k, b_{n-k})$$

est alors absolument convergente, de somme

$$B\left(\left(\sum_{n=0}^{+\infty} a_n\right), \left(\sum_{n=0}^{+\infty} b_n\right)\right)$$

La démonstration est, pour l'essentiel, analogue à celle qui précède.

Terminons par un exercice, donnant un résultat de "sommation par paquets", qui s'appliquera notamment aux séries complexes absolument convergentes :

EXERCICE 9-5.27 Si $(u_i)_{i\in I}$ est une famille sommable et $J \subset I$ une partie (infinie) de I, montrer que $(u_i)_{i\in J}$ est sommable. On peut donc donner un sens au symbole

$$\sum_{i\in J} u_i$$

ce symbole ayant aussi un sens si J est finie.

Montrer que, si $I = \bigcup_{n\in\mathbb{N}} J_n$ est une **partition** de I en une suite de parties finies ou dénombrables, la série de terme général

$$S_n = \sum_{i\in J_n} u_i$$

est absolument convergente et

$$\sum_{i\in I} u_i = \sum_{k=0}^{+\infty} \left(\sum_{i\in J_k} u_i \right)$$

9-6 La fonction exponentielle

Dans cette section, nous définirons les fonctions exponentielle et circulaires comme somme de séries, ce qui permettra de résoudre avec précision le problème de la mesure des angles.

9-6.1 Fonction exponentielle complexe

9-6.1.1 Définition. Equation fonctionnelle

DÉFINITION 9-6.1 *Pour $z \in \mathbb{C}$, on pose*

$$\exp z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

série absolument convergente (par exemple par application de la règle de D'Alembert).

THÉORÈME 9-6.2 La fonction $\exp : (\mathbb{C}, +) \to (\mathbb{C}^*, \times)$ **est un morphisme de groupes.**

Démonstration : Le terme général du produit de Cauchy des séries absolument convergentes définissant $\exp z_1$ et $\exp z_2$ pour $z_1, z_2 \in \mathbb{C}$ est

$$c_n = \sum_{k=0}^{n} \frac{z_1^k}{k!} \frac{z_2^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} C_n^k z_1^k z_2^{n-k} = \frac{(z_1 + z_2)^n}{n!}$$

par la formule du binôme. Le théorème **9-5**.24 donne alors

$$\exp(z_1 + z_2) = \exp z_1 \exp z_2$$

En particulier, pour $z_2 = -z_1$, on obtient

$$\exp z_1 \exp(-z_1) = \exp 0 = 1$$

ce qui montre que la fonction exponentielle ne s'annule jamais. ■

DÉFINITION 9-6.3 *On définit le réel e par*

$$e = \exp 1 = \sum_{n=0}^{+\infty} \frac{1}{n!}$$

EXERCICE 9-6.4 Montrer que e est irrationnel. Indication : si S_n est somme partielle de la série définissant e, on pourra d'abord montrer que

$$S_n < e < S_n + \frac{1}{nn!}$$

et arriver à une contradiction en supposant e rationnel.

On montre facilement en utilisant la propriété de morphisme de groupes que, pour $m \in \mathbb{Z}$

$$\exp(m) = e^m$$

Comme la fonction exponentielle prend des valeurs strictement positives sur $\mathbb{R}$ (car $\exp(x) = \left(\exp \frac{x}{2}\right)^2$), ce résultat est généralisable au cas où m est un rationnel : si $m = \dfrac{p}{q}$ avec $q \in \mathbb{N}^*$, on a

$$(\exp m)^q = \exp mq = \exp p = e^p$$

ce qui donne bien (puisque $\exp m > 0$)

$$\exp m = (e^p)^{\frac{1}{q}} = e^{\frac{p}{q}} = e^m$$

Il n'y a alors pas d'ambiguïté à **définir** le symbole e^z par

$$\forall z \in \mathbb{C} \quad e^z = \exp z$$

9-6.1.2 Module d'une exponentielle

THÉORÈME 9-6.5 Pour $y \in \mathbb{R}$, e^{iy} est un complexe de module 1. Pour $z \in \mathbb{C}$

$$|e^z| = e^{\operatorname{Re} z}$$

Démonstration : On a en effet

$$\overline{e^{iy}} = \overline{\sum_{k=0}^{+\infty} \frac{(iy)^k}{k!}} = \sum_{k=0}^{+\infty} \frac{(-iy)^k}{k!} = e^{-iy} = \frac{1}{e^{iy}}$$

puisque $y \in \mathbb{R}$. On obtient bien

$$\left|e^{iy}\right|^2 = e^{iy}\overline{e^{iy}} = 1$$

Pour $z = x + iy \in \mathbb{C}$, on a enfin

$$|e^z| = |e^x| \left|e^{iy}\right| = |e^x| = e^x$$

puisque e^x est un réel strictement positif. ■

A ce stade, rappelons que le module d'un complexe est tout simplement $|a+ib| = \sqrt{a^2+b^2}$, et vérifie $|zz'| = |z|\,|z'|$. Nous n'avons pas pour le moment de définition de l'argument d'un complexe non nul, définition qui utilisera justement la fonction exponentielle complexe.

9-6.1.3 Propriété de continuité et de dérivabilité

Nous verrons dans un chapitre ultérieur (séries de fonctions) des théorèmes généraux qui permettraient de prouver les résultats de cette section. Nous donnerons ici des démonstrations "à la main", utilisant pour l'essentiel l'équation fonctionnelle vérifiée par l'exponentielle.

THÉORÈME 9-6.6 La fonction $\exp : \mathbb{C} \to \mathbb{C}^*$ est continue.

Démonstration : Pour z et $h \in \mathbb{C}$, on a

$$\left|e^{z+h} - e^z\right| = e^{\operatorname{Re} z}\left|e^h - 1\right|$$

ce qui prouve que la continuité de l'exponentielle en z sera conséquence de la continuité en 0. On a de plus

$$e^h - 1 = \sum_{k=1}^{+\infty} \frac{h^k}{k!}$$

série absolument convergente, ce qui donne pour $|h| < 1$

$$\left|e^h - 1\right| \leqslant |h| \sum_{k=1}^{+\infty} \frac{1}{k!}$$

et prouve la continuité de la fonction exponentielle en 0. ■

Si on se restreint au domaine réel, on a le caractère C^∞ de la fonction exponentielle :

THÉORÈME 9-6.7 La fonction $f = \exp|_{\mathbb{R}}$ est C^∞, avec $f' = f$ sur $\mathbb{R}$. Elle réalise une bijection strictement croissante de $\mathbb{R}$ dans $\mathbb{R}^{*+}$. On a

$$\forall k \in \mathbb{N} \qquad \begin{cases} x^k \underset{x \to +\infty}{=} o\left(e^x\right) \\ e^x \underset{x \to -\infty}{=} o\left(\dfrac{1}{x^k}\right) \end{cases}$$

Démonstration : On étudie cette fois, pour x et $h \in \mathbb{R}$ la quantité

$$\Delta(h) = e^{x+h} - e^x - he^x = e^x\left(e^h - 1 - h\right) = e^x \sum_{k=2}^{+\infty} \frac{h^k}{k!}$$

Pour $|h| < 1$, on a clairement

$$|\Delta(h)| \leqslant |h|^2 e^x \sum_{k=2}^{+\infty} \frac{1}{k!}$$

ce qui donne $\Delta(h) = O(h^2)$ quand $h \to 0$, et prouve la dérivabilité de la fonction exponentielle en x, la dérivée valant e^x en ce point. Pour $x > 0$ et $k \in \mathbb{N}$, on a

$$e^x \geqslant \frac{x^{k+1}}{(k+1)!}$$

ce qui donne $\lim_{+\infty} x^{-k}e^x = +\infty$ et donc $x^k = o(e^x)$ en $+\infty$. L'étude en $-\infty$ en résulte, puisque $e^x = \frac{1}{e^{-x}}$. La dérivée de la fonction exponentielle étant partout > 0, cette fonction est strictement croissante. Tendant vers 0 en $-\infty$ et vers $+\infty$ en $+\infty$, elle réalise donc un C^∞-difféomorphisme de $\mathbb{R}$ dans $]0, +\infty[$. ■

On pourrait à partir de là, et sans recourir à l'intégration, définir la fonction logarithme $\mathbb{R}^{*+} \to \mathbb{R}$ comme étant le difféomorphisme réciproque de l'exponentielle. C'est aussi un isomorphisme du groupe multiplicatif des réels > 0 vers le groupe additif réel. On vérifie aisément que

$$\forall x > 0 \quad (\ln)'(x) = \frac{1}{x}$$

DÉFINITION 9-6.8 *Si x est un réel strictement positif, et $\alpha \in \mathbb{C}$, on pose*

$$x^\alpha = e^{\alpha \ln x}$$

cette définition est cohérente avec la définition de x^ρ lorsque $x > 0$ et ρ est un rationnel.

REMARQUE 9-6.9 Des techniques de majoration analogues à celles utilisées dans le cas réel démontreraient, pour z et h **variables complexes**

$$e^{z+h} = e^z + he^z + O\left(h^2\right) \text{ pour } h \to 0$$

ce qui prouve en particulier que

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{C}^*}} \frac{e^{z+h} - e^z}{h} = e^z$$

On dit alors que la fonction exponentielle est dérivable (comme fonction d'une variable complexe, voir à ce sujet l'exercice **11-3**.41 et la section **23-1.4**) en tout point de $\mathbb{C}$, avec

$$(\exp)'(z) = e^z$$

EXERCICE 9-6.10 Si $\alpha \in \mathbb{C}$, montrer que $f :]0, +\infty[\to \mathbb{C}$ définie par $x \mapsto x^\alpha$ est dérivable et que

$$\forall x > 0 \quad f'(x) = \alpha x^{\alpha - 1}$$

9-6.2 Les fonctions circulaires

9-6.2.1 Définitions. Dérivabilité

DÉFINITION 9-6.11 *Pour $t \in \mathbb{R}$, on définit*

$$\cos t = \operatorname{Re} e^{it} \quad \textit{et} \quad \sin t = \operatorname{Im} e^{it}$$

On a donc $\forall t \in \mathbb{R} \quad e^{it} = \cos t + i \sin t$, et la propriété $|e^{it}| = 1$ se traduit par

$$\forall t \in \mathbb{R} \quad \cos^2 t + \sin^2 t = 1$$

Toutes les formules de trigonométrie circulaire (cosinus d'une somme, d'une différence etc...) se déduisent aisément de la formule $\forall t, t' \in \mathbb{R} \quad e^{i(t+t')} = e^{it}e^{it'}$. La continuité de la fonction exponentielle donne immédiatement

THÉORÈME 9-6.12 Les fonctions cos **et** sin **sont continues de** $\mathbb{R}$ **dans** $[-1, 1]$.

On a d'ailleurs clairement, à l'aide de la remarque **9-6**.9

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{R}^*}} \frac{e^{i(t+h)} - e^{it}}{ih} = e^{it}$$

En séparant les parties réelles et imaginaires, on obtient

THÉORÈME 9-6.13 Les fonctions cos **et** sin **sont** C^{∞} **sur** $\mathbb{R}$ **et**

$$\forall t \in \mathbb{R} \quad (\cos)'(t) = -\sin t \quad \text{et} \quad (\sin)'(t) = \cos t$$

Pour $t \in \mathbf{R}$, le développement

$$e^{it} = \sum_{k=0}^{+\infty} \frac{(it)^k}{k!}$$

donne, en séparant les parties réelles et imaginaires :

$$\forall t \in \mathbb{R} \quad \cos t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} t^{2n} \quad \text{et} \quad \sin t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1}$$

séries toutes deux absolument convergentes.

9-6.2.2 Périodicité. Le nombre π

Il n'est pas évident, sur les développements en séries précédents, que les fonctions sin et cos prennent toutes leurs valeurs dans $[-1, 1]$. C'est l'équation fonctionnelle vérifiée par la fonction exponentielle qui a donné ce résultat, par l'intermédiaire de

$$|e^{it}| = 1 \Rightarrow \cos^2 t + \sin^2 t = 1$$

De même, les propriétés de périodicité des fonctions circulaires vont apparaître grâce à cette même équation fonctionnelle. Commençons par découvrir le nombre π :

PROPOSITION 9-6.14 La fonction cos **possède un plus petit zéro dans** $]0, +\infty[$. **Si** α **désigne ce plus petit zéro, on pose**

$$\pi = 2\alpha$$

Démonstration : On montre d'abord que la fonction cos s'annule au moins une fois dans l'intervalle $[0, 2]$. Comme il s'agit d'une fonction continue, cela résultera du théorème des valeurs intermédiaires, si l'on vérifie que $\cos 2 < 0$ (puisque $\cos 0 = 1$). Or

$$\cos 2 = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} 4^k$$

est somme d'une série alternée. En notant v_k la valeur absolue du terme général, on a

$$\frac{v_{k+1}}{v_k} = \frac{2}{(k+1)(2k+1)}$$

ce qui montre que cette valeur absolue commence à décroître à partir du rang $k = 1$. On a donc l'encadrement

$$S_1 \leqslant \cos 2 \leqslant S_2$$

avec $S_2 = 1 - 2 + \frac{2}{3} = -\frac{1}{3} < 0$. On considère ensuite

$$\alpha = \inf \{t \in [0, 2] \mid \cos t = 0\}$$

On a $\alpha > 0$ car la fonction cos est strictement positive au voisinage de 0. De plus, par définition de la borne inf, α peut être approché par la droite par des zéros de la fonction cos, et par continuité de celle-ci on obtient $\cos \alpha = 0$. Par construction, α est la plus petite solution dans $\mathbb{R}^+$ de l'équation $\cos t = 0$. ∎

COROLLAIRE 9-6.15 On a

$$e^{i\frac{\pi}{2}} = i \quad \textbf{et} \quad e^{2i\pi} = 1$$

Démonstration : On a $\cos \dfrac{\pi}{2} = 0$ et donc $\sin \dfrac{\pi}{2} = \pm 1$. Comme on sait que la dérivée de la fonction sin est strictement positive sur $\left[0, \dfrac{\pi}{2}\right[$, cette fonction est croissante sur cet intervalle, et comme $\sin 0 = 0$ on a forcément $\sin \dfrac{\pi}{2} = 1$ et donc

$$e^{i\frac{\pi}{2}} = i$$

On en déduit $e^{i\pi} = \left(e^{i\frac{\pi}{2}}\right)^2 = -1$ et $e^{2i\pi} = 1$. ∎

On peut en déduire le noyau du morphisme de groupes

$$\exp : (\mathbb{C}, +) \to (\mathbb{C}^*, \times)$$

COROLLAIRE 9-6.16 Pour $z \in \mathbb{C}$, on a

$$e^z = 1 \Leftrightarrow z \in 2i\pi\mathbb{Z}$$

Démonstration : On a évidemment, par propriété de morphisme

$$z \in 2i\pi\mathbb{Z} \Rightarrow e^z = 1$$

Réciproquement, si $z = x + iy$ est solution de cette équation, on a

$$|e^z| = e^x = 1 \Rightarrow x = 0$$

puisqu'on sait que la fonction exponentielle est injective en restriction à $\mathbb{R}$. Il reste donc à résoudre

$$e^{iy} = 1$$

avec $y \in \mathbb{R}$. L'ensemble des solutions est un sous-groupe additif de $\mathbb{R}$. Pour montrer que c'est $2\pi\mathbb{Z}$, il suffit de voir que la plus petite solution strictement positive est 2π, c'est-à-dire qu'il n'y a pas de solution dans $]0, 2\pi[$. S'il existait une telle solution, on aurait

$$0 < \frac{y}{4} < \frac{\pi}{2}$$

et donc $u = \cos \frac{y}{4} > 0$ et $v = \sin \frac{y}{4} > 0$ (définition de π et croissance du sinus sur $\left]0, \frac{\pi}{2}\right[$). On aurait alors

$$e^{iy} = (u + iv)^4 = u^4 - 6u^2v^2 + v^4 + 4iuv\left(u^2 - v^2\right) = 1$$

En regardant la partie imaginaire on obtient $u^2 = v^2 = \frac{1}{2}$ (puisque $u^2 + v^2 = 1$). Mais alors

$$e^{iy} = (u + iv)^4 = u^4 - 6u^2v^2 + v^4 = -1$$

ce qui est contraire aux hypothèses. ■

Comme $e^z = e^{z'} \Leftrightarrow e^{z-z'} = 1$, on obtient

$$\boxed{\forall z, z' \in \mathbb{C} \quad e^z = e^{z'} \Leftrightarrow z - z' \in 2i\pi\mathbb{Z}}$$

COROLLAIRE 9-6.17 Les fonctions $\sin$ et $\cos$ sont 2π-périodiques (2π est la plus petite période > 0).

Démonstration : La 2π-périodicité est conséquence de ce qui précède. Si T est une période du cosinus, c'est une période de sa dérivée. Les deux fonctions sont donc T périodiques et il en est de même de la fonction $t \mapsto e^{it}$. T est donc un multiple de 2π. ■

Le tableau de variation

t	0		$\frac{\pi}{2}$
$\cos t$	1	$\searrow$	0
$\sin t$	0	$\nearrow$	1

peut être facilement étendu à tout l'intervalle $[0, 2\pi]$ par

$$\sin\left(t + \frac{\pi}{2}\right) = \cos t, \ \cos\left(t + \frac{\pi}{2}\right) = -\sin t$$

9-6.2.3 Mesure des angles. Argument d'un complexe

THÉORÈME 9-6.18 L'application $t \mapsto e^{it}$ est un morphisme surjectif de $(\mathbb{R}, +) \to (\mathbb{U}, \times)$ groupe des complexes de module 1. Son noyau est $2\pi\mathbb{Z}$.

Démonstration : Il suffit de démontrer que tout complexe de module 1 peut s'écrire sous la forme $e^{i\theta}$, avec $\theta \in \mathbb{R}$. Comme le noyau du morphisme considéré est exactement $2\pi\mathbb{Z}$, le réel θ sera unique modulo 2π. Soit $\omega = a + ib \in \mathbb{U}$. Si $a \geqslant 0$ et $b \geqslant 0$, on peut trouver un $\theta \in \left[0, \frac{\pi}{2}\right]$ avec $\cos\theta = a$ (théorème des valeurs intermédiaires). On a alors

$$\sin^2\theta = 1 - \cos^2\theta = 1 - a^2 = b^2$$

et comme $\sin\theta \geqslant 0$, on a $\sin\theta = b$, soit finalement $e^{i\theta} = \omega$. Si $a < 0$ et $b \geqslant 0$, le complexe $-i\omega \in \mathbb{U}$ a ses parties réelle et imaginaire positives, et s'écrit donc $e^{i\alpha}$, pour $\alpha \in \mathbb{R}$. On a alors

$$-i\omega = e^{i\alpha} \Rightarrow \omega = e^{i\left(\alpha + \frac{\pi}{2}\right)}$$

Enfin, si $b < 0$, en travaillant avec $-\omega$ dont la partie imaginaire est positive, on obtiendra une écriture

$$-\omega = e^{i\alpha} \Rightarrow \omega = e^{i(\alpha+\pi)}$$

ce qui prouve la surjectivité de l'application considérée. ■

DÉFINITION 9-6.19 *Si z est un complexe non nul, on appelle argument de z la classe de réels modulo 2π définie par*

$$\operatorname{Arg} z = \left\{\theta \in \mathbb{R} \mid e^{i\theta} = \frac{z}{|z|}\right\}$$

Tout réel appartenant à cette classe sera dit détermination de l'argument de z. On notera souvent

$$\theta = \arg z \quad \textit{pour} \quad \theta \in \operatorname{Arg} z$$

sachant bien qu'il ne s'agit pas d'une véritable égalité.

On a évidemment, pour z et z' complexes non nuls

$$\operatorname{Arg}(zz') = \operatorname{Arg} z + \operatorname{Arg} z' \text{ et } \operatorname{Arg}\frac{z'}{z} = \operatorname{Arg} z' - \operatorname{Arg} z$$

L'unique détermination de l'argument de $z \in \mathbb{C}^*$ qui appartient à l'intervalle $]-\pi, \pi]$ est souvent appelée **détermination principale de l'argument de z.** On définit ainsi une application $\mathbb{C}^* \to]-\pi, \pi]$ non continue en tout point de l'axe réel négatif, mais **continue en restriction à $\mathbb{C} - \mathbb{R}^-$** :

Si $z = x + iy \in \mathbb{C} - \mathbb{R}^-$ et si $\theta \in]-\pi, \pi[$ est détermination principale de son argument, on a

$$\tan\frac{\theta}{2} = \frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}} = \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{2\cos^2\frac{\theta}{2}} = \frac{\sin\theta}{1+\cos\theta} = \frac{\frac{y}{\sqrt{x^2+y^2}}}{1+\frac{x}{\sqrt{x^2+y^2}}} = \frac{y}{x+\sqrt{x^2+y^2}}$$

Comme $\frac{\theta}{2}$ tombe dans le bon intervalle, on a

$$\boxed{\theta = 2\arctan\frac{y}{x+\sqrt{x^2+y^2}}}$$

9-6.2.4 Résolution de $e^z = a$

THÉORÈME 9-6.20 Si $a \in \mathbb{C}^*$, l'équation $e^z = a$ possèdent une infinité de solutions :

$$e^z = a \Leftrightarrow z = \ln|a| + i\operatorname{Arg} a$$

Une solution de cette équation est appelée détermination du logarithme complexe de a. Celle correspondant au choix de la détermination principale de l'argument de a est appelée détermination principale du logarithme (complexe) de a. On la notera $\operatorname{Ln} a$.

Démonstration : Exercice.

EXERCICE 9-6.21 Montrer que $a \mapsto \operatorname{Ln} a$ est continue sur $\mathbb{C} - \mathbb{R}^-$. En utilisant la remarque 9-6.9, montrer que cette fonction de la variable complexe a est dérivable en tout point de $\mathbb{C} - \mathbb{R}^-$ et que

$$(\operatorname{Ln})'(a) = \frac{1}{a}$$

9-6.2.5 Dérivée d'une exponentielle complexe

EXERCICE 9-6.22 Montrer que, si I est un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{C}$ est dérivable, il en est de même de $g : I \to \mathbb{C}$ définie par

$$g(t) = e^{f(t)}$$

et que

$$\forall t \in I \quad g'(t) = f'(t)\, e^{f(t)}$$

9-6.3 Exponentielle dans une algèbre de Banach

Dans tout ce qui suit $(A, +, \times, ., \| \|)$ est une algèbre de Banach, c'est-à-dire une algèbre normée complète avec

$$\forall x, y \in A \quad \|xy\| \leqslant \|x\|\,\|y\|$$

Dans la pratique, ce sera pour nous une algèbre de matrices carrées $\mathcal{M}_p(\mathbb{K})$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, munie d'une norme subordonnée à une norme quelconque sur $\mathbb{K}^p$. On pourra aussi envisager le cas $A = \mathcal{L}_c(\mathbb{E})$, algèbre des endomorphismes continus d'un espace de Banach.

9-6.3.1 Fonction exponentielle

THÉORÈME 9-6.23 (ET DEFINITION) **Si** $(A,+,\times,.,\|\,\|)$ **est une algèbre de Banach et** $a \in A$**, la série de terme général** $\frac{a^n}{n!}$ **est absolument convergente. Sa somme est par définition l'exponentielle de** a :

$$\exp a = \sum_{k=0}^{+\infty} \frac{a^k}{k!}$$

(par convention, le premier terme vaut $a^0 = 1_A$**).**

Démonstration : Il suffit de remarquer que

$$\left\|\frac{a^k}{k!}\right\| \leqslant \frac{\|a\|^k}{k!}$$

terme général d'une série convergente. ■

REMARQUE 9-6.24 Si la série converge pour une norme, elle converge pour toute norme équivalente. Le résultat précédent subsiste dans toute algèbre A munie d'une norme N pour laquelle A est complète et le produit continu. On sait alors qu'il existe une constante $M > 0$ avec $\forall x, y \quad N(xy) \leqslant MN(x)\,N(y)$. La norme $\|x\| = MN(x)$ vérifie alors $\|xy\| \leqslant \|x\|\,\|y\|$.

Par exemple, pour $\lambda \in \mathbb{K}$

$$\exp(\lambda 1_A) = e^{\lambda} 1_A$$

La remarque **9-5**.26 montre que l'on peut, dans le cas d'une algèbre de Banach, refaire la démonstration du théorème **9-6**.2, pour peu que l'on puisse encore appliquer la formule du binôme. On a donc

THÉORÈME 9-6.25 Si a **et** b **sont deux éléments qui commutent dans une algèbre de Banach, alors**

$$\exp(a+b) = \exp(a)\exp(b)$$

Le résultat ne subsiste plus si $ab \neq ba$, comme le montre l'exercice suivant :

EXERCICE 9-6.26 Dans $\mathcal{M}_2(\mathbb{R})$, on pose

$$a = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Calculer $\exp(a)$, $\exp(b)$ et $\exp(a+b)$.

COROLLAIRE 9-6.27 Si $a \in A$ **et** $t, s \in \mathbb{K}$**, on a**

$$\exp(ta)\exp(sa) = \exp((s+t)\,a)$$

En particulier, l'égalité

$$\exp(a)\exp(-a) = 1_A$$

montre que, dans une algèbre de Banach, une exponentielle est toujours inversible.

9-6.3.2 Cas particulier : exponentielle d'une matrice

Si l'on doit calculer explicitement l'exponentielle d'une matrice $M \in \mathcal{M}_n(\mathbb{K})$, on peut d'abord penser simplifier M par similitude, puisque si $P \in \mathcal{GL}_n(\mathbb{K})$

$$M = PNP^{-1} \Rightarrow \forall m \in \mathbb{N} \quad \sum_{k=0}^{m} \frac{M^k}{k!} = P\left(\sum_{k=0}^{m} \frac{N^k}{k!}\right) P^{-1}$$

ce qui, en faisant tendre m vers $+\infty$, compte tenu de la continuité du produit matriciel donne

$$M = PNP^{-1} \Rightarrow \exp M = P \exp(N) P^{-1}$$

Cette formule n'est pas étonnante. Il y a derrière cette égalité la notion d'exponentielle d'un endomorphisme d'un $\mathbb{K}$-ev de dimension n. Si $\mathbf{E}$ est un tel espace, le choix d'une norme arbitraire sur $\mathbb{E}$ (elles sont toutes équivalentes) munit $\mathcal{L}(\mathbf{E}) = \mathcal{L}_c(\mathbf{E})$ de la norme des applications linéaires continues, donc d'une structure d'algèbre de Banach. Pour $u \in \mathcal{L}(\mathbf{E})$, on a

$$\exp u = \sum_{k=0}^{+\infty} \frac{u^k}{k!}$$

et si $\mathcal{B}$ est une base de $\mathbf{E}$, avec $M = M(u, \mathcal{B})$, on aura

$$M(\exp u, \mathcal{B}) = \exp M$$

Un cas simple est évidemment **le cas où M est diagonalisable**. Il est en effet clair que

$$D = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n) \Rightarrow \exp D = \mathrm{Diag}(e^{\alpha_1}, \ldots, e^{\alpha_n})$$

et donc

$$M = P\,\mathrm{Diag}(\alpha_1, \ldots, \alpha_n) P^{-1} \Rightarrow \exp M = P\,\mathrm{Diag}(e^{\alpha_1}, \ldots, e^{\alpha_n}) P^{-1}$$

Le cas où M ne possède qu'une seule valeur propre complexe (et n'est pas scalaire) est en fait encore plus simple, bien que dans ce cas M ne soit pas diagonalisable : on sait dans ce cas que, si λ est cette valeur propre,

$$M = \lambda I_n + N$$

avec N nilpotente d'indice inférieur à n. Comme I_n et M commutent, on a alors

$$\exp M = e^{\lambda} \exp N$$

et le calcul de $\exp N$ est simple, puisque la série qui définit cette matrice a tous ses termes nuls à partir d'un certain rang. On a finalement

$$\exp M = e^{\lambda} \left(\sum_{k=0}^{n-1} \frac{N^k}{k!}\right)$$

Dans le cas général, on se ramènera à la situation précédente, en déterminant une base de chacun des sous-espaces caractéristiques de l'endomorphisme de $\mathbb{C}^n$ canoniquement associé à M (on fera les calculs dans $\mathcal{M}_n(\mathbb{C})$). On obtiendra ainsi une matrice de passage P telle que

$$M = P\begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_p \end{pmatrix} P^{-1}$$

où les A_i sont des matrices carrées n'ayant qu'une valeurs propre. On aura alors

$$\exp M = P\begin{pmatrix} \exp A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \exp A_p \end{pmatrix} P^{-1}$$

EXERCICE 9-6.28 Montrer que, pour $M \in \mathcal{M}_n(\mathbb{K})$

$$\det \exp M = e^{\operatorname{trace} M}$$

Cette notion d'exponentielle intervient lors de la résolution de systèmes différentiels ("à coefficients constants"). Nous y reviendrons en détail ultérieurement. Contentons nous de quelques résultats, caractérisant l'exponentielle par une équation différentielle.

9-6.3.3 Equation différentielle $x' = u(x)$

THÉORÈME 9-6.29 Si a est élément d'une algèbre de Banach A, l'application $\varphi : \mathbb{R} \to A$ définie par $\varphi(t) = \exp(ta)$ est dérivable sur $\mathbb{R}$ et vérifie

$$\forall t \in \mathbb{R} \quad \varphi'(t) = a\exp(ta) = \exp(ta)\,a$$

Démonstration : Etudions la différence

$$\Delta(h) = \exp((t+h)a) - \exp(ta) - h\exp(ta)\,a$$

On a $\Delta(h) = \exp(ta)(\exp(ha) - 1_A - ha)$, ce qui donne

$$\|\Delta(h)\| \leqslant \|\exp(ta)\| \sum_{k=2}^{+\infty} \frac{|h|^k \|a\|^k}{k!}$$

ce qui donne facilement

$$\|\Delta(h)\| = O(h^2)$$

et prouve le résultat. ■

COROLLAIRE 9-6.30 Soit $\mathbb{E}$ un espace de Banach et $u \in \mathcal{L}_c(\mathbb{E})$ un endomorphisme continu de $\mathbb{E}$. Pour tout $x_0 \in \mathbb{E}$, l'équation différentielle

$$x'(t) = u(x(t))$$

possède une unique solution définie sur $\mathbb{R}$ vérifiant la condition initiale $x(0) = x_0$. Cette solution est donnée par

$$x(t) = \exp(tu)(x_0)$$

Démonstration : Rappelons qu'une solution définie sur un intervalle I de $\mathbb{R}$ est une fonction dérivable $x : I \to \mathbb{E}$ vérifiant

$$\forall t \in I \quad x'(t) = u(x(t))$$

La fonction $x(t) = \exp(tu)(x_0)$ est bien définie sur $\mathbb{R}$. Elle est dérivable car composée de la fonction dérivable $t \mapsto \exp(tu)$ (théorème précédent appliqué à l'algèbre de Banach $\mathcal{L}_c(\mathbb{E})$) de $\mathbb{R}$ dans $\mathcal{L}_c(\mathbb{E})$ et de l'application linéaire continue $\mathcal{L}_c(\mathbb{E}) \to \mathbb{E}$ définie par $v \mapsto v(x_0)$. On a bien

$$x(0) = x_0 \text{ et } x'(t) = u\exp(tu)(x_0) = u(x(t))$$

ce qui prouve que cette fonction est bien solution de l'équation différentielle pour la condition initiale donnée. Pour montrer que c'est la seule, on recopie la démonstration qu'on connaît bien dans le cas $\mathbb{E} = \mathbb{R}$: on fait "varier une constante". Si $t \mapsto y(t)$ est une autre solution, on considère

$$z(t) = \exp(-tu)(y(t))$$

z est clairement dérivable sur $\mathbb{R}$ (on utilise cette fois la continuité de l'application bilinéaire $\mathcal{L}_c(\mathbb{E}) \times \mathbb{E} \to \mathbb{E}$, définie par $(v, x) \mapsto v(x)$). On a

$$\begin{aligned} z'(t) = -u(z(t)) + \exp(-tu)(y'(t)) &= -u(z(t)) + \exp(-tu)\,u(y(t)) \\ &= -u(z(t)) + u(z(t)) = 0_{\mathbb{E}} \end{aligned}$$

puisque y est solution de l'équation différentielle et que u et $\exp(tu)$ commutent. On en déduit que z est constante égale à $z(0) = y(0) = x_0$, ce qui donne aisément

$$y(t) = \exp(tu)(x_0)$$

et prouve l'unicité de la solution pour une condition initiale donnée.
■

Si on avait imposé la condition initiale $x(t_0) = x_0$, la solution s'écrirait alors

$$x(t) = \exp((t - t_0)u)(x_0)$$

Le théorème précédent s'applique notamment aux systèmes d'équations linéaires à coefficients constants s'écrivant

$$\left\{ \begin{array}{lclcl} x_1'(t) & = & a_{11}x_1(t) & + \cdots + & a_{1n}x_n(t) \\ \vdots & & & \vdots & \\ x_i'(t) & = & a_{i1}x_1(t) & + \cdots + & a_{in}x_n(t) \\ \vdots & & & \vdots & \\ x_n'(t) & = & a_{n1}x_1(t) & + \cdots + & a_{nn}x_n(t) \end{array} \right.$$

où les x_i sont des fonctions inconnues $\mathbb{R} \to \mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. En considérant le vecteur

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \in \mathbb{K}^n$$

et la matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, le système s'écrira

$$X'(t) = AX(t)$$

(ou, si l'on préfère $X' = \Phi_A(X)$), et la solution pour la condition initiale $X(t_0) = X_0 \in \mathbb{K}^n$ sera

$$X(t) = \exp((t - t_0) A) X_0$$

9-7 Exercices

EXERCICE 9-7.1 Soit a_n une suite de réels décroissante vers 0.

1. Si $\sum a_n$ converge, montrer que $na_n \to 0$. La réciproque est-elle vraie ?
2. Montrer que $\sum a_n$ et $\sum n(a_{n-1} - a_n)$ sont de même nature.
3. $\sum a_n$, $\sum n a_{n^2}$ et $\sum 2^n a_{2^n}$ sont de même nature (critère de "condensation"). Qu'obtient-on pour les séries de Bertrand ?

EXERCICE 9-7.2 Pour $p \geqslant 0$ et $u_0 \in \mathbb{R}$ on définit la suite u_n par : $u_n = \dfrac{\sin u_{n-1}}{n^p}$.
Nature de la série de terme général u_n ?
Même question avec $u_{n+1} = \dfrac{1}{n} e^{-u_n}$. Nature de $\sum (-1)^n u_n$?

EXERCICE 9-7.3 Dans la suite $\dfrac{1}{n}$ on supprime un terme, on garde le suivant, on supprime 2 termes, on garde le suivant, on supprime 3 termes, on garde le suivant etc... Quelle est la nature de la série ainsi obtenue ?
Dans la série harmonique, on supprime tous les termes dont le dénominateur a une écriture décimale contenant au moins un 9. Quelle est la nature de cette série ?

EXERCICE 9-7.4 Soit $u_0 > 0$ et soit $(a_n)_{n\geqslant 0}$ une suite de $\mathbb{R}^+$. On définit alors la suite $(u_n)_{n\geqslant 0}$ par la récurrence $u_{n+1} = u_n + (a_n/u_n)$. Prouver que la suite (u_n) converge ssi la série de terme général a_n converge.

EXERCICE 9-7.5 Soit $0 < \alpha < 1$, $\sum u_n$ et $\sum v_n$ deux séries à termes strictement positifs convergentes. Montrer que la série de terme général $w_n = u_n^\alpha v_n^{1-\alpha}$ est convergente et que sa somme vérifie

$$\sum w_n \leqslant \left(\sum u_n\right)^\alpha \left(\sum v_n\right)^{1-\alpha}$$

Généraliser à un nombre quelconque de séries.

EXERCICE 9-7.6 Soit $\sum u_n$ une série à termes > 0 divergente et $S_n = \sum\limits_0^n u_k$. Nature des séries de termes généraux $\dfrac{u_n}{S_n}$ et $\dfrac{u_n}{S_n^\alpha}$?

EXERCICE 9-7.7 Donner un développement asymptotique contenant :

1. 3 termes pour $\sum_{k=1}^{n} \frac{1}{\sqrt{k}}$.

2. 3 termes infiniment petits non nuls pour $\sum_{k=1}^{n} \frac{\ln k}{k}$.

EXERCICE 9-7.8 Quelle est la partie entière de $\sum_{1}^{999} \frac{1}{k^{\frac{2}{3}}}$?

EXERCICE 9-7.9 Soit $x \in \mathbb{R}$ et $u_n = \frac{1}{C_{n+2}^2} + \frac{x}{C_{n+1}^2} + \cdots + \frac{x^p}{C_{n+2-p}^2} + \cdots + \frac{x^n}{C_2^2}$. Nature et somme de la série de terme général u_n ?

EXERCICE 9-7.10 Pour $x \in]0,1[$ et $\beta \in]0,1[$ montrer que $\beta(1-x) < 1 - x^\beta$. Soit $\sum a_n$ une série convergente à termes positifs, $r_n = \sum_{k=n}^{+\infty} a_k$ et $\rho \in]0,1[$. Montrer que la série $\sum \frac{a_n}{r_n^\rho}$ converge et que

$$\sum_{n=1}^{+\infty} \frac{a_n}{r_n^\rho} \leqslant \frac{1}{1-\rho}\Big(\sum_{n=1}^{+\infty} a_n\Big)^{1-\rho}$$

EXERCICE 9-7.11 Soit S l'ensemble des suites $u = (u_n)_{n\in\mathbb{N}}$ de réels positifs avec $\sum_{n\in\mathbb{N}} u_n = 1$. Déterminer

$$\inf_{u\in S} \sum_{n=0}^{+\infty} \Big[u_n\big(\sum_{k=0}^{n} u_k\big)\Big]$$

EXERCICE 9-7.12 Soit $(a_n)_{n\geqslant 0}$ une suite réelle croissante telle que

$$a_0 = 1 \text{ et } \lim_{n\to+\infty} a_n = +\infty$$

On considère la suite (u_n) définie par $u_n = \left(1 - \frac{a_{n-1}}{a_n}\right) \frac{1}{\sqrt{a_n}}$, $n \geqslant 1$.

1. Prouver que la série de terme général u_n converge et que sa somme est majorée par 2.

2. Prouver enfin que pour tout $\alpha \in]0,2[$, il existe une suite (a_n) telle que $\sum_{n=1}^{\infty} u_n = \alpha$.

EXERCICE 9-7.13 Soit $\mathcal{P}$ la parabole d'équation $y^2 = 2px$ dans un repère orthonormé et $M_0 \in \mathcal{P}$. La normale en M_0 à $\mathcal{P}$ recoupe $\mathcal{P}$ en M_1. La normale en M_1 à $\mathcal{P}$ recoupe $\mathcal{P}$ en M_2 etc... On construit ainsi une suite de points $M_n(x_n, y_n) \in \mathcal{P}$. Etudier la nature de la série $\sum \frac{1}{y_n}$.

EXERCICE 9-7.14 Convergence de $\sum_{n=1}^{+\infty} e^{-\sqrt{\ln n}}$ et $\sum_{n=1}^{+\infty} e^{-\sqrt{n}}$.

EXERCICE 9-7.15 Nature de la série de terme général $u_n = \dfrac{S_n}{n^\alpha}$ où $S_n = \sum_{p=2}^{n} (\ln p)^3$.

EXERCICE 9-7.16 On note $f(n) = \sum_{p=0}^{n-1} \dfrac{2}{2n+2p+1}$. Quelle est la nature des séries de termes généraux $u_n = n^{-f(n)}$ et $v_n = (-1)^n u_n$?

EXERCICE 9-7.17 On pose $u_n = \dfrac{(-1)^{n-1}}{n^\alpha} = (\alpha > 0)$ et on note R_n le reste d'ordre n de la série $\sum u_n$. Etudier la convergence de $\sum R_n$.

EXERCICE 9-7.18 On définit la suite u_n par $u_0 = 1$ et $u_{n+1} = (n+a+b)u_n$, où a et b sont deux réels > 0. Quelle est la nature de la série de terme général

$$v_n = (-1)^n \frac{u_n}{n!} \frac{2n+a+b}{(n+a)(n+b)}$$

EXERCICE 9-7.19 Soit $\alpha > 0$ fixé. Donner un équivalent simple de $S_n = \sum_{p=1}^{n} p^{n\alpha}$ pour $n \to +\infty$.

EXERCICE 9-7.20 Quelques notions sur les produits infinis :
Définition : Soit $(z_n)_{n\in\mathbb{N}}$ une suite complexe ne s'annulant pas. On dit que le produit infini $\prod_{n=0}^{+\infty} z_n$ converge ssi la suite des produits partiels $P_n = \prod_{k=0}^{n} z_k$ admet une limite non nulle. Dans les autres cas, on dit que le produit infini est divergent.

1) a) Montrer que si le produit infini est convergent on a $\lim_{n\to+\infty} z_n = 1$. On posera donc $z_n = 1 + u_n$ avec $u_n \to 0$. On suppose dans un premier temps que $u_n \in \mathbb{R}$.

b) Si on suppose que $u_n > -1$ à partir d'un certain rang, montrer que le produit infini $\prod_{0}^{+\infty}(1+u_n)$ est convergent ssi la série $\sum_{n=0}^{+\infty} \ln(1+u_n)$ est convergente. En déduire que si de plus u_n garde un signe constant, le produit converge ssi la série $\sum u_n$ converge.

c) Si u_n ne garde pas un signe constant et si $\sum u_n$ converge, montrer que le produit infini et la série $\sum u_n^2$ sont de même nature.

d) Etudier le cas $u_n = \dfrac{(-1)^n}{\sqrt{n}} + \dfrac{1}{2n}$.

2) On suppose maintenant $u_n \in \mathbb{C}$.

a) Démontrer que pour $\alpha_1, \cdots, \alpha_n \in \mathbb{C}$ on a

$$\left|\prod_{i=1}^{n}(1+\alpha_i) - 1\right| \leqslant \left(\prod_{i=1}^{n}(1+|\alpha_i|) - 1\right)$$

b) On dit que le produit $\prod(1+u_n)$ est absolument convergent ssi la série $\sum |u_n|$ converge. Montrer que tout produit absolument convergent est convergent.

c) On suppose $\sum u_n$ convergente et $\sum u_n^2$ absolument convergente. Montrer que le produit $\prod(1+u_n)$ converge. On pourra étudier le produit infini $\prod(1+u_n)e^{-u_n}$ après avoir montré que, pour $z \in \mathbb{C}$, on a

$$|e^z - 1 - z| \leqslant \frac{|z|^2}{2} e^{|z|}$$

EXERCICE 9-7.21 Soit $\alpha \in \mathbb{R}$. Nature de la série $\displaystyle\sum_{n=1}^{+\infty} \frac{(-1)^{E(\sqrt{n})}}{n^\alpha}$.

EXERCICE 9-7.22 Nature de la série $\displaystyle\sum_{n>|\alpha|}^{+\infty} \ln\left(\frac{\sqrt{n}+(-1)^n}{\sqrt{n+\alpha}}\right)$ où $\alpha \in \mathbb{R}$.

Chapitre 10

Intégration sur un segment

10-1 Intégrale d'une fonction continue par morceaux

10-1.1 Intégrale d'une fonction en escalier

10-1.1.1 Définition

Dans tout ce chapitre $(E, \| \|)$ représente un espace vectoriel normé complet. Ce sera le plus souvent $E = \mathbb{R}$ ou $\mathbb{C}$, ou un espace vectoriel de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. S'il est de dimension finie égale à n, nous le noterons E_n. Conformément à la définition **8-2**.1, nous considérons l'espace $\mathcal{E}([a,b], E)$ des fonctions en escaliers définies sur un segment $[a,b]$ à valeurs dans E. Cet espace est muni de la norme de la convergence uniforme

$$\|s\|_\infty = \sup_{t\in[a,b]} \|s(t)\|$$

DÉFINITION 10-1.1 *Soit $s \in \mathcal{E}([a,b], E)$ et $d : a = t_0 < t_1 < t_2 < \cdots < t_p = b$ une subdivision de $[a,b]$ adaptée à s (ce qui signifie que s est constante égale à $x_i \in E$ sur chaque intervalle ouvert $]t_i, t_{i+1}[$ défini par la subdivision). On définit l'intégrale de s sur $[a,b]$ comme étant le vecteur de E*

$$\int_{[a,b]} s = \sum_{i=0}^{p-1} (t_{i+1} - t_i)\, x_i$$

Ce vecteur est indépendant de la subdivision adaptée à s. Cette intégrale est aussi notée $\int_a^b s(t)\,dt$

Si on intercale un point de subdivision t' entre t_i et t_{i+1}, on ne change pas l'intégrale puisque

$$(t_{i+1} - t_i)\,x_i = (t_{i+1} - t')\,x_i + (t' - t_i)\,x_i$$

Il en est de même en ajoutant un nombre quelconque (fini !) de points de subdivision. Si d' est une autre subdivision de $[a, b]$ adaptée à s, si on note $I(d)$ la valeur de l'intégrale calculée avec d, on aura alors $I(d) = I(d \cup d') = I(d')$ puisque d et d' jouent des rôles symétriques.

Dans l'écriture $\int_a^b s(t)\,dt$, la lettre t est variable muette : l'intégrale est un vecteur de E qui ne dépend pas d'une variable t ! Cette notation trouvera sa justification en permettant des changements de variables "automatiques" comme nous le verrons plus loin.

10-1.1.2 Propriétés fondamentales

Les propriétés suivantes découlent aisément de la définition :

- **On ne change pas l'intégrale en modifiant la fonction en un nombre fini de points de** $[a, b]$.

  Modifier un fonction s en un nombre fini de points revient en effet simplement à rajouter des points à une subdivision adaptée à s.

- **Linéarité.**

  L'application $\mathcal{E}([a, b], E) \to E$ définie par

  $$s \mapsto \int_a^b s(t)\,dt$$

  est une application linéaire. L'homogénéité est évidente. Pour montrer que l'intégrale d'une somme $s_1 + s_2$ est somme des intégrales de s_1 et s_2, on utilisera évidemment une subdivision de $[a, b]$ adaptée à la fois à s_1 et s_2.

  Plus généralement, si $(F, \|\ \|)$ est un autre espace de Banach et si on se donne $u \in \mathcal{L}_c(E, F)$ une **application linéaire** continue[1] de E vers F, on a

  $$\forall s \in \mathcal{E}([a, b], E) \quad u\left(\int_a^b s(t)\,dt\right) = \int_a^b u(s(t))\,dt$$

  égalité qui est conséquence immédiate de la linéarité de u.

[1] Pour le moment, le fait que E et F soient complets ou u soit continue n'a pas d'importance. Ces hypothèses deviendront indispensables lorsqu'on intègrera des fonctions autres que des fonctions en escalier.

- **Cas de la dimension finie.**

Dans la pratique, nous travaillerons le plus souvent avec un espace E_n de dimension finie. Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est une base de E_n, une fonction en escalier $s \in \mathcal{E}([a,b], E)$ s'écrira

$$s = \sum_{i=1}^{n} s_i \, e_i$$

avec $s_i \in \mathcal{E}([a,b], \mathbb{K})$. L'expression de l'intégrale de s dans la base $\mathcal{B}$ sera alors évidemment

$$\int_a^b s(t)\, dt = \sum_{i=1}^{n} \left(\int_a^b s_i(t)\, dt \right) e_i$$

- **Inégalité de la norme.**

Si $s \in \mathcal{E}([a,b], E)$, l'application $t \to \|s(t)\|$ est évidemment dans $\mathcal{E}([a,b], \mathbb{R})$, et l'inégalité triangulaire donne immédiatement

$$\left\| \int_a^b s(t)\, dt \right\| \leqslant \int_a^b \|s(t)\|\, dt$$

Ceci vaut évidemment pour les fonctions réelles ou complexes, où la norme est simplement la valeur absolue ou le module.

- **Continuité par rapport à la norme de la convergence uniforme.**

L'application linéaire $s \mapsto \int_a^b s(t)\, dt$ est continue de $(\mathcal{E}([a,b], E), \| \|_\infty)$ dans $(E, \| \|)$, puisqu'une majoration évidente dans l'inégalité précédente donnera

$$\left\| \int_a^b s(t)\, dt \right\| \leqslant (b-a)\, \|s\|_\infty$$

- **Additivité par rapport à l'intervalle d'intégration.**

Si $s \in \mathcal{E}([a,b], E)$ et c est un réel vérifiant $a < c < b$, il est évident que $s|_{[a,c]} \in \mathcal{E}([a,c], E)$ et $s|_{[c,b]} \in \mathcal{E}([c,b], E)$. Par un abus de notation commode (on ne change pas le nom d'une fonction quand on restreint son domaine de définition), nous noterons $\int_{[a,c]} s = \int_a^c s(t)\, dt$ l'intégrale de la restriction de s à $[a,c]$. En ajoutant le point c à une subdivision adaptée à s, on obtient facilement

$$\int_a^b s(t)\, dt = \int_a^c s(t)\, dt + \int_c^b s(t)\, dt$$

10-1.1.3 Cas des fonctions réelles

L'espace $\mathcal{E}([a,b],\mathbb{R})$ est muni d'une relation d'ordre **partiel**[2] :

$$s_1 \leqslant s_2 \Leftrightarrow \forall t \in [a,b] \quad s_1(t) \leqslant s_2(t)$$

PROPOSITION 10-1.2 L'intégrale est une forme linéaire positive[3] sur l'espace $\mathcal{E}([a,b],\mathbb{R})$ c'est-à-dire que

$$\forall s \in \mathcal{E}([a,b],\mathbb{R}) \quad 0 \leqslant s \Rightarrow 0 \leqslant \int_a^b s(t)\,dt$$

Ceci résulte directement de la définition de l'intégrale. Comme pour s_1 et $s_2 \in \mathcal{E}([a,b],\mathbb{R})$, on a

$$s_1 \leqslant s_2 \Leftrightarrow 0 \leqslant s_2 - s_1$$

par linéarité de l'intégrale, on a la possibilité d'intégrer des inégalités :

COROLLAIRE 10-1.3 Pour s_1 et $s_2 \in \mathcal{E}([a,b],\mathbb{R})$

$$s_1 \leqslant s_2 \Leftrightarrow \int_a^b s_1(t)\,dt \leqslant \int_a^b s_2(t)\,dt$$

En particulier, comme on a toujours

$$s \in \mathcal{E}([a,b],\mathbb{R}) \Rightarrow \left(\inf_{[a,b]} s\right) 1_{[a,b]} \leqslant s \leqslant \left(\sup_{[a,b]} s\right) 1_{[a,b]}$$

on aura

$$\left(\inf_{[a,b]} s\right)(b-a) \leqslant \int_a^b s(t)\,dt \leqslant \left(\sup_{[a,b]} s\right)(b-a)$$

Remarquons pour conclure que l'on peut évidemment intégrer des inégalités entre fonctions de $\mathcal{E}([a,b],\mathbb{R})$ qui seraient valables sur $[a,b]$ privé d'un nombre fini de points.

10-1.2 Intégrale d'une fonction continue par morceaux

D'après le corollaire **8-2**.5, toute fonction f continue par morceaux de $[a,b]$ dans E est limite uniforme d'une suite de fonctions en escalier. La continuité de l'intégrale définie sur $\mathcal{E}([a,b],E)$ par rapport à la norme de la convergence uniforme va permettre d'utiliser une technique déjà décrite à la section **7-2.4** pour prolonger l'intégrale à $C_m^0([a,b],E)$.

[2] La négation de $s_1 \leqslant s_2$ n'est pas $s_1 > s_2$!

[3] Ne pas se méprendre sur cette terminologie : une forme linéaire non identiquement nulle sur un $\mathbb{R}$-espace vectoriel quelconque est surjective ... et ne peut donc prendre toutes ses valeurs dans $\mathbb{R}^+$.

10-1.2.1 Définition

THÉORÈME 10-1.4 (ET DEFINITION) Si $f \in \mathcal{C}_m^0([a,b],E)$, pour toute suite $(s_n)_{n\in\mathbb{N}}$ de fonctions en escalier convergeant uniformément vers f sur $[a,b]$ la suite des intégrales $\left(\int_a^b s_n(t)\,dt\right)_{n\in\mathbb{N}}$ est convergente dans $(E,\|\ \|)$, et sa limite ne dépend pas de la suite $(s_n)_{n\in\mathbb{N}}$ choisie. Cette limite est appelée intégrale de f sur $[a,b]$, et on la note

$$\int_a^b f(t)\,dt = \lim_{n\to+\infty}\int_a^b s_n(t)\,dt$$

Démonstration : Comme $(E,\|\ \|)$ est complet, il suffit de prouver que la suite $\left(\int_a^b s_n(t)\,dt\right)_{n\in\mathbb{N}}$ est de Cauchy. Par inégalité de la norme, on a

$$\left\|\int_a^b s_n(t)\,dt - \int_a^b s_m(t)\,dt\right\| \leqslant (b-a)\,\|s_n - s_m\|_\infty$$

Cette quantité peut être rendue inférieure à un réel $\varepsilon > 0$ arbitraire pour n et m suffisamment grands, puisque la suite $(s_n)_{n\in\mathbb{N}}$ étant convergente dans $(\mathcal{B}([a,b],E),\|\ \|_\infty)$ est évidemment de Cauchy dans cet espace. En "mélangeant" deux suites quelconques $(s_n^1)_{n\in\mathbb{N}}$ et $(s_n^2)_{n\in\mathbb{N}}$ de fonctions en escalier convergeant vers une fonction f, c'est à dire en posant

$$s_{2n} = s_n^1 \text{ et } s_{2n+1} = s_n^2$$

on montre facilement que la limite des intégrales ne dépend pas de la suite choisie. Lorsque f est elle-même en escalier, le choix de la suite constante (égale à f) montre montre que l'intégrale sur $\mathcal{C}_m^0([a,b],E)$ prolonge celle sur $\mathcal{E}([a,b],E)$ ■

Ce passage à la limite conserve les propriétés de l'intégrale sur l'espace des fonctions en escalier :

10-1.2.2 Propriétés fondamentales

- **On ne change pas l'intégrale en modifiant la fonction en un nombre fini de points de $[a,b]$.**

  Si f est modifiée sur un ensemble fini $X \subset [a,b]$ de points, on modifie sur les mêmes points les fonctions de la suite s_n, en imposant par exemple

$$\forall x\in X \quad \forall n\in\mathbb{N} \quad s_n(x) = f(x)$$

  Cette opération conserve évidemment la convergence uniforme, et ne modifie ni les intégrales des s_n, ni leur limite.

- **Linéarité.**

L'application $\mathcal{C}_m^0([a,b],E) \to E$ définie par

$$f \mapsto \int_a^b f(t)\, dt$$

est une application linéaire. En effet, si f et $g \in \mathcal{C}_m^0([a,b],E)$, et si $(s_n)_{n\in\mathbb{N}}$ et $(t_n)_{n\in\mathbb{N}}$ sont deux suites de fonctions en escalier convergeant respectivement vers f et g, pour α appartenant au corps de base ($\mathbb{R}$ ou $\mathbb{C}$), on a

$$\lim_{n\to+\infty} \|(f+\alpha g)-(s_n+\alpha t_n)\|_\infty = 0$$

et il suffit de passer à la limite dans les égalités

$$\forall n \in \mathbb{N} \quad \int_{[a,b]} (s_n+\alpha t_n) = \int_{[a,b]} s_n + \alpha \int_{[a,b]} t_n$$

pour obtenir

$$\int_{[a,b]} (f+\alpha g) = \int_{[a,b]} f + \alpha \int_{[a,b]} g$$

Plus généralement, si $(F, \| \|)$ est un autre espace de Banach et si u est une **application linéaire continue**[4] de E vers F, on a

$$\forall f \in \mathcal{C}_m^0([a,b],E) \quad u\left(\int_a^b f(t)\, dt\right) = \int_a^b u(f(t))\, dt$$

En effet, la fonction $u \circ f$ est dans $\mathcal{C}_m^0([a,b],F)$, et si $(s_n)_{n\in\mathbb{N}}$ est une suite de fonctions de $\mathcal{E}([a,b],E)$ qui converge uniformément vers f, on a évidemment $u \circ s_n \in \mathcal{E}([a,b],F)$ avec

$$\forall n \in \mathbb{N} \quad \|u\circ f - u \circ s_n\|_\infty \leqslant \|u\|\, \|f - s_n\|_\infty$$

(où $\|u\|$ est la norme de l'application linéaire continue u). La suite de fonctions en escalier $(u\circ s_n)_{n\in\mathbb{N}}$ converge donc uniformément vers $u \circ f$, et l'égalité

$$u\left(\int_a^b f(t)\, dt\right) = \int_a^b u(f(t))\, dt$$

s'obtient ici encore par passage à la limite.

- **Cas de la dimension finie.**

Si $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ est une base de E_n, et si $f \in \mathcal{C}_m^0([a,b],E)$, on aura

$$f = \sum_{i=1}^n f_i\, e_i$$

[4] Ici, la continuité de u est importante !

avec $f_i \in \mathcal{C}_m^0([a,b],\mathbb{K})$. L'expression de l'intégrale de f dans la base $\mathcal{B}$ sera alors évidemment

$$\int_a^b f(t)\,dt = \sum_{i=1}^n \left(\int_a^b f_i(t)\,dt\right) e_i$$

En particulier, lorsque l'on considère $\mathbb{C}$ comme $\mathbb{R}$-ev de dimension 2, rapporté à sa base canonique $\{1, i\}$, on obtient

$$\forall f \in \mathcal{C}_m^0([a,b],\mathbb{C}) \quad \int_a^b f(t)\,dt = \int_a^b \operatorname{Re} f(t)\,dt + i\int_a^b \operatorname{Im} f(t)\,dt$$

et en conséquence

$$\forall f \in \mathcal{C}_m^0([a,b],\mathbb{C}) \quad \overline{\int_a^b f(t)\,dt} = \int_a^b \overline{f(t)}\,dt$$

- **Inégalité de la norme.**

  Si $f \in \mathcal{C}_m^0([a,b],E)$, l'application $t \to \|f(t)\|$ est dans $\mathcal{C}_m^0([a,b],\mathbb{R})$, et on a

$$\left\|\int_a^b f(t)\,dt\right\| \leqslant \int_a^b \|f(t)\|\,dt$$

  Pour les fonctions réelles ou complexes, ceci donnera

$$\left|\int_a^b f(t)\,dt\right| \leqslant \int_a^b |f(t)|\,dt$$

  Pour obtenir ces inégalités, ici encore, on passera à la limite, pour une suite $(s_n)_{n\in\mathbb{N}}$ de fonctions en escalier convergeant uniformément vers f, dans les inégalités

$$\left\|\int_a^b s_n(t)\,dt\right\| \leqslant \int_a^b \|s_n(t)\|\,dt$$

  en utilisant d'une part la continuité de la norme et d'autre part le fait que

$$\forall t \in [a,b] \quad \big|\|f(t)\| - \|s_n(t)\|\big| \leqslant \|f(t) - s_n(t)\| \leqslant \|f - s_n\|_\infty$$

  ce qui prouve la convergence uniforme de la suite de fonctions en escalier $t \mapsto \|s_n(t)\|$ vers la fonction continue par morceaux $t \mapsto \|f(t)\|$.

- **Continuité par rapport à la norme de la convergence uniforme.**

  L'application linéaire $f \mapsto \displaystyle\int_a^b f(t)\,dt$ est continue de $(\mathcal{C}_m^0([a,b],E), \|\ \|_\infty)$ dans $(E, \|\ \|)$, puisqu'on a encore ici

$$\left\|\int_a^b f(t)\,dt\right\| \leqslant (b-a)\,\|f\|_\infty$$

  inégalité qu'on obtient aussi par passage à la limite : on a en effet

$$\big|\|f\|_\infty - \|s_n\|_\infty\big| \leqslant \|f - s_n\|_\infty$$

  Ce résultat est suffisamment important pour être énoncé sous forme de théorème :

THÉORÈME 10-1.5 Si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions continues par morceaux de $[a,b]$ dans E qui converge uniformément sur $[a,b]$ vers une fonction f continue par morceaux, alors

$$\lim_{n\to+\infty}\int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt = \int_a^b \lim_{n\to+\infty} f_n(t)\,dt$$

On obtient ici une **condition suffisante**[5] **pour "permuter passage à la limite et symbole d'intégration"**. Par contre, il importe de noter que la convergence simple n'est pas une condition suffisante pour aboutir à la même conclusion. En effet, sur $[0,1]$, la suite de fonctions en escalier

$$s_n(t) = n.1_{]0,\frac{1}{n}[}(t)$$

converge simplement vers la fonction nulle, et pourtant la suite des intégrales est constante égale à 1, et ne tend pas vers 0.

Insistons lourdement : la démonstration du théorème précédent se ramène à l'écriture de l'inégalité

$$\left\|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right\| \leqslant \int_a^b \|f(t) - f_n(t)\|\,dt \leqslant (b-a)\,\|f_n - f\|_\infty$$

et la longueur de l'intervalle d'intégration, soit $(b-a)$, intervient dans cette majoration. Lorsque, dans un chapitre ultérieur, nous intègrerons sur des intervalles non bornés, on conçoit que la convergence uniforme ne sera alors plus suffisante pour permuter les symboles de limite et intégration.

- **Additivité par rapport à l'intervalle d'intégration.**

  Avec les mêmes abus de notation que dans le cas des fonctions en escalier, nous aurons, pour $f \in C_m^0([a,b],E)$ et $c \in]a,b[$

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$$

  égalité qui s'obtient ici encore par passage à la limite.

REMARQUE 10-1.6 Le procédé de construction utilisé ici permettrait en fait de prolonger l'intégrale à toute fonction de $[a,b]$ dans E qui serait limite uniforme d'une suite de fonctions en escalier, c'est-à-dire à toute fonction réglée sur $[a,b]$ (*cf.* exercice **8-2**.6). Les propriétés qui précèdent subsisteraient.

[5] Cette condition n'est pas nécessaire. Par exemple, sur $[0,1]$, la suite de fonctions définie par $f_n(t) = t^n$ converge simplement, mais pas uniformément vers la fonction $g = 1_{\{1\}}$, fonction caractéristique du singleton $\{1\}$. On a cependant

$$\lim_{n\to+\infty}\int_0^1 t^n\,dt = 0 = \int_0^1 g(t)\,dt$$

10-1.2.3 Intégration et inégalités

Comme dans le cas des fonctions en escalier, le résultat fondamental est :

PROPOSITION 10-1.7 L'intégrale est une forme linéaire positive sur l'espace vectoriel $C_m^0([a,b],\mathbb{R})$.

Démonstration : Si $f \in C_m^0([a,b],\mathbb{R})$ est positive et $(s_n)_{n\in\mathbb{N}}$ est une suite de fonctions en escalier qui converge uniformément vers f, la majoration

$$|f - |s_n|| \leqslant |f - s_n|$$

montre que la suite $(|s_n|)_{n\in\mathbb{N}}$ convetge aussi uniformément vers f, et par conséquent, puisque les intégrales de ces fonctions sont positives

$$\int_a^b f(t)\,dt = \lim_{n\to+\infty} \int_a^b |s_n(t)|\,dt \geqslant 0 \quad \blacksquare$$

COROLLAIRE 10-1.8 Si f et $g \in C_m^0([a,b],\mathbb{R})$ vérifient $f \leqslant g$ sur $[a,b]$ (sauf peut-être en un nombre fini de points) alors

$$\int_a^b f(t)\,dt \leqslant \int_a^b g(t)\,dt$$

COROLLAIRE 10-1.9 Si $f \in C_m^0([a,b],\mathbb{R})$, on a

$$\inf_{[a,b]} f.1_{[a,b]} \leqslant f \leqslant \sup_{[a,b]} f.1_{[a,b]}$$

et donc

$$\inf_{[a,b]} f.(b-a) \leqslant \int_a^b f(t)\,dt \leqslant \sup_{[a,b]} f.(b-a)$$

COROLLAIRE 10-1.10 Si $f \in C_m^0([a,b],\mathbb{R})$ est positive alors

$$\int_a^b f(t)\,dt = 0 \Rightarrow f \text{ est nulle en tout point de continuité}$$

En d'autres termes, puisque f est supposée continue par morceaux, la fonction f est nulle sur $[a,b]$, sauf peut-être en un nombre fini de points. En particulier, si f est continue sur $[a,b]$, elle est identiquement nulle.

Démonstration : Soit t_0 un point de continuité de f, que nous supposerons intérieur à $[a,b]$ (le cas d'une extrémité se traite de façon analogue). Supposons $f(t_0) > 0$. On peut, par continuité, trouver un réel $\alpha > 0$ tel que

$$[t_0-\alpha, t_0+\alpha] \subset [a,b] \quad \text{et}$$
$$\forall t \in [t_0-\alpha, t_0+\alpha] \quad |f(t) - f(t_0)| \leqslant \frac{f(t_0)}{2}$$

On a alors pour $t \in [t_0 - \alpha, t_0 + \alpha]$

$$f(t) \geqslant \frac{f(t_0)}{2}$$

et, par additivité par rapport à l'intervalle d'intégration

$$\int_a^b f(t)\,dt = \int_a^{t_0-\alpha} f(t)\,dt + \int_{t_0-\alpha}^{t_0+\alpha} f(t)\,dt + \int_{t_0+\alpha}^b f(t)\,dt$$

ce qui donne finalement, puisque f est positive

$$\int_a^b f(t)\,dt \geqslant \int_{t_0-\alpha}^{t_0+\alpha} f(t)\,dt \geqslant \alpha f(t_0) > 0$$

ce qui contredit l'hypothèse. ■

EXERCICE 10-1.11 Trouver toutes les fonctions continues $f : [a,b] \to \mathbb{C}$ vérifiant

$$\left|\int_a^b f(t)\,dt\right| = \int_a^b |f(t)|\,dt$$

On étudiera d'abord le cas d'une fonction réelle. Lorsque f est complexe, on posera $\int_a^b f(t)\,dt = \rho e^{i\theta}$ avec $\rho \in \mathbb{R}^+$ et $\theta \in \mathbb{R}$, et on considèrera la fonction $g = e^{-i\theta} f$.

REMARQUE 10-1.12 Il arrive souvent que l'on ait à majorer le module d'une intégrale d'un produit de deux fonctions continues par morceaux, $f, g : [a,b] \to \mathbb{C}$. Nous verrons dans la section suivante que l'inégalité de Schwarz permet d'obtenir un majorant. Parfois, c'est une intégration par parties qui donnera une autre majoration. Par l'inégalité de la norme, on a aussi

$$\left|\int_a^b f(t)\,g(t)\,dt\right| \leqslant \int_a^b |f(t)|\,.\,|g(t)|\,dt$$

ce qui donne par exemple l'inégalité (puisque $|f|\,.\,|g| \leqslant \|f\|_\infty\,.\,|g|$)

$$\left|\int_a^b f(t)\,g(t)\,dt\right| \leqslant \|f\|_\infty \int_a^b |g(t)|\,dt$$

Par contre majorer[6] par $\|f\|_\infty \left|\int_a^b g(t)\,dt\right|$ est une lourde faute (sauf situation particulière, par exemple si g est réelle de signe constant). Qu'obtiendrait-on par exemple si l'intégrale de g était nulle ? Méditer en particulier le cas $g(t) = \sin t$ sur $[a,b] = [0, 2\pi]$, avec par exemple $f = g$.

[6]Cette erreur provient d'une mauvaise manipulation d'inégalités : si f et g sont réelles, on ne peut passer de $-\|f\|_\infty \leqslant f \leqslant \|f\|_\infty$ à $-\|f\|_\infty\, g \leqslant fg \leqslant \|f\|_\infty\, g$ que lorsque g est positive !

10-1.2.4 Inégalités de Schwarz et Minkowski

THÉORÈME 10-1.13 (INEGALITE DE SCHWARZ) Si f et g sont continues par morceaux de $[a,b]$ dans $\mathbb{C}$, on a

$$\left|\int_a^b f(t)\overline{g(t)}\,dt\right|^2 \leqslant \int_a^b |f(t)|^2\,dt.\int_a^b |g(t)|^2\,dt$$

Si f et g sont de plus continues, il y a égalité si et seulement si f et g sont linéairement dépendantes.

Démonstration : L'inégalité est évidente (et est une égalité) si l'une des intégrales du second membre est nulle : si, par exemple, $\int_a^b |f(t)|^2\,dt = 0$, la fonction f est nulle sur $[a,b]$, sauf peut-être en un nombre fini de points, et il en est de même de $f\overline{g}$. De même, si le premier membre est nul, l'inégalité est claire. Si les deux membres sont non nuls, évaluons, pour $\lambda \in \mathbb{C}$

$$\int_a^b |\lambda f(t)+g(t)|^2\,dt = |\lambda|^2\int_a^b |f(t)|^2\,dt + 2\,\mathrm{Re}\left(\lambda\int_a^b f(t)\overline{g(t)}\,dt\right) + \int_a^b |g(t)|^2\,dt$$

Si

$$\int_a^b f(t)\overline{g(t)}\,dt = \rho e^{i\theta}$$

avec $\rho > 0$, on obtient, pour $\lambda = xe^{-i\theta}$ avec x réel quelconque

$$\forall x \in \mathbb{R} \qquad x^2\int_a^b |f(t)|^2\,dt + 2\rho x + \int_a^b |g(t)|^2\,dt \geqslant 0$$

Le discriminant de ce polynôme du second degré est donc négatif ou nul, ce qui donne

$$\rho^2 - \int_a^b |f(t)|^2\,dt.\int_a^b |g(t)|^2\,dt \leqslant 0$$

ce qui est exactement l'inégalité à démontrer.

Supposons à présent l'égalité avec f et g continues. Si les deux membres sont nuls, on a par exemple $\int_a^b |f(t)|^2\,dt = 0$, ce qui donne $f \equiv 0$, et donc $\{f,g\}$ liée. Si f et g sont non nulles, le trinôme du second degré étudié plus haut a un discriminant nul, donc possède une racine (double). En particulier

$$\exists\lambda \in \mathbb{C} \quad \int_a^b |\lambda f(t)+g(t)|^2\,dt = 0$$

et, comme la fonction intégrée est continue positive, $\lambda f + g = 0$, ce qui donne encore $\{f, g\}$ liée.

Enfin, si $\{f, g\}$ est liée (avec f et g non nécessairement continues), il est clair que l'inégalité de Schwarz est un égalité (prendre par exemple $g = \alpha f$, avec $\alpha \in \mathbb{C}$). ■

COROLLAIRE 10-1.14 (INEGALITE DE MINKOWSKI) Si f et g sont continues par morceaux de $[a, b]$ dans $\mathbb{C}$, on a

$$\left(\int_a^b |f(t) + g(t)|^2 \, dt\right)^{\frac{1}{2}} \leqslant \left(\int_a^b |f(t)|^2 \, dt\right)^{\frac{1}{2}} + \left(\int_a^b |g(t)|^2 \, dt\right)^{\frac{1}{2}}$$

Lorsque f et g sont continues, il y a égalité si et seulement si f est nulle ou $(\exists \alpha \in \mathbb{R}^+ \quad g = \alpha f)$.

Démonstration : L'inégalité à démontrer portant sur des réels positifs est équivalente à

$$\int_a^b |f(t) + g(t)|^2 \, dt \leqslant \left[\left(\int_a^b |f(t)|^2 \, dt\right)^{\frac{1}{2}} + \left(\int_a^b |g(t)|^2 \, dt\right)^{\frac{1}{2}}\right]^2$$

ce qui équivaut à

$$\mathrm{Re}\left(\int_a^b f(t)\,\overline{g(t)}\, dt\right) \leqslant \left(\int_a^b |f(t)|^2 \, dt\right)^{\frac{1}{2}} \cdot \left(\int_a^b |g(t)|^2 \, dt\right)^{\frac{1}{2}}$$

Cette inégalité est conséquence de l'inégalité de Schwarz, puisque la partie réelle d'un complexe est toujours inférieure à son module.

Si f et g sont continues, il est clair que si elles sont $\mathbb{R}^+$-proportionnelles, l'inégalité de Minkowski est en fait une égalité. Réciproquement, pour avoir l'égalité avec f et g continues, il faut déjà avoir égalité dans l'inégalité de Schwarz, ce qui impose

$$f = 0 \quad \text{ou} \quad \exists \alpha \in \mathbb{C} \quad g = \alpha f$$

Il faut de plus avoir

$$\mathrm{Re}\left(\int_a^b f(t)\,\overline{g(t)}\, dt\right) = \left|\int_a^b f(t)\,\overline{g(t)}\, dt\right|$$

ce qui donne, dans le cas où f n'est pas nulle, avec $g = \alpha f$

$$\mathrm{Re}\left(\overline{\alpha}\int_a^b |f(t)|^2 \, dt\right) = |\alpha| \int_a^b |f(t)|^2 \, dt$$

soit $\overline{\alpha} = |\alpha|$ et donc $\alpha \in \mathbb{R}^+$. ■

REMARQUE 10-1.15 Les raisonnements qui précèdent sont un cas particulier d'un résultat plus général que nous verrons dans un chapitre d'algèbre bilinéaire : les inégalités précédentes proviennent du fait que, sur $\mathcal{C}_m^0([a,b],\mathbb{C})\times\mathcal{C}_m^0([a,b],\mathbb{C})$, l'application

$$(f,g)\mapsto\langle f,g\rangle=\int_a^b\overline{f(t)}g(t)\,dt$$

est une forme sesquilinéaire hermitienne telle que la forme quadratique associée soit positive :

$$\forall f\in\mathcal{C}_m^0([a,b],\mathbb{C})\quad\langle f,f\rangle=\int_a^b|f(t)|^2\,dt\geqslant 0$$

Cette forme, en restriction à $\mathcal{C}^0([a,b],\mathbb{C})$, est définie positive et munit cet espace d'une structure d'espace préhilbertien complexe. C'est pour cela que, dans l'inégalité de Schwarz, une des fonctions est "conjuguée". Comme g et $\overline{g}$ ont même modules, on pourrait aussi écrire cette inégalité sous la forme

$$\left|\int_a^b f(t)\,g(t)\,dt\right|^2\leqslant\int_a^b|f(t)|^2\,dt.\int_a^b|g(t)|^2\,dt$$

Dans le cas de fonctions réelles, on travaillerait avec la forme bilinéaire symétrique

$$\mathcal{C}_m^0([a,b],\mathbb{R})\times\mathcal{C}_m^0([a,b],\mathbb{R})\to\mathbb{R}\qquad(f,g)\mapsto\langle f,g\rangle=\int_a^b f(t)\,g(t)\,dt$$

la forme quadratique associée étant positive. L'espace

$$\left(\mathcal{C}^0([a,b],\mathbb{R}),\langle\rangle\right)$$

est un espace préhilbertien réel. L'inégalité de Schwarz peut alors s'écrire

$$\left(\int_a^b f(t)\,g(t)\,dt\right)^2\leqslant\int_a^b f^2(t)\,dt.\int_a^b g^2(t)\,dt$$

(il importe évidemment d'écrire les modules dans le cas de fonctions complexes, pour ne pas écrire d'inégalités entre nombres ... complexes !)

EXERCICE 10-1.16 (INEGALITE DE HÖLDER) Deux réels p et $q\in]1,+\infty[$ sont dits **conjugués** si et seulement si

$$\frac{1}{p}+\frac{1}{q}=1$$

Montrer que, si p et q sont conjugués, pour α et $\beta\geqslant 0$, on a

$$\alpha\beta\leqslant\frac{\alpha^p}{p}+\frac{\beta^q}{q}$$

Montrer que, si f et g sont des fonctions continues par morceaux de $[a,b]\to\mathbb{C}$, on a, pour p et q conjugués

$$\left|\int_a^b f(t)\,g(t)\,dt\right|\leqslant\left(\int_a^b|f(t)|^p\,dt\right)^{\frac{1}{p}}.\left(\int_a^b|g(t)|^q\,dt\right)^{\frac{1}{q}}$$

(Inégalité de Hölder). Indication : se ramener d'abord à travailler avec des fonctions réelles positives. L'inégalité est évidente si le second membre est nul. Si ce n'est pas le cas, poser

$$\alpha(t) = \frac{f(t)}{\left(\int_a^b |f(t)|^p \, dt\right)^{\frac{1}{p}}} \quad \text{et} \quad \beta(t) = \frac{g(t)}{\left(\int_a^b |g(t)|^q \, dt\right)^{\frac{1}{q}}}$$

et utiliser la majoration de $\alpha\beta$ précédente pour conclure.

On peut déduire de l'inégalité de Hölder l'inégalité de Minkowski (généralisée) : pour

$$f \text{ et } g \in \mathcal{C}_m^0([a,b], \mathbb{C})$$

on a

$$\forall p > 1 \qquad \left(\int_a^b |f(t) + g(t)|^p \, dt\right)^{\frac{1}{p}} \leqslant \left(\int_a^b |f(t)|^p \, dt\right)^{\frac{1}{p}} + \left(\int_a^b |g(t)|^p \, dt\right)^{\frac{1}{p}}$$

(Indication : se ramener au cas où f et g sont réelles positives, poser

$$q = \frac{p}{p-1}$$

Ecrire

$$(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}$$

et majorer l'intégrale de chacun des produits au second membre en utilisant l'inégalité de Hölder).

10-1.2.5 Convergence en moyenne et en moyenne quadratique

Pour $f \in \mathcal{C}_m^0([a,b], \mathbb{C})$, on pose

$$\|f\|_1 = \int_a^b |f(t)| \, dt \quad \text{et} \quad \|f\|_2 = \left(\int_a^b |f(t)|^2 \, dt\right)^{\frac{1}{2}}$$

PROPOSITION 10-1.17 Les applications $f \mapsto \|f\|_1$ et $f \mapsto \|f\|_2$ définissent des semi-normes sur l'espace $\mathcal{C}_m^0([a,b], \mathbb{C})$. On les appelle respectivement semi-norme de la convergence en moyenne sur $[a,b]$ et semi-norme de la convergence en moyenne quadratique sur $[a,b]$. En restriction à l'espace $\mathcal{C}^0([a,b], \mathbb{C})$ des fonctions continues sur $[a,b]$, ce sont des normes.

Démonstration : L'homogénéité est claire. L'inégalité triangulaire est évidente pour $\| \|_1$, à cause de la majoration $|f+g| \leqslant |f| + |g|$. Pour $\| \|_2$, c'est l'inégalité de Minkowski. Enfin, d'après le corollaire **10-1**.10,

$$\forall f \in \mathcal{C}_m^0([a,b], \mathbb{C}) \quad \|f\|_1 = 0 \Leftrightarrow \|f\|_2 = 0$$
$$\Leftrightarrow f \text{ est nulle sauf en un nombre fini de points.}$$

ce qui montre que $\| \|_1$ et $\| \|_2$ ne sont pas véritablement des normes sur $\mathcal{C}_m^0([a,b], \mathbb{C})$, mais le sont en restriction à $\mathcal{C}^0([a,b], \mathbb{C})$. ■

DÉFINITION 10-1.18 *Soient $(f_n)_{n\in\mathbb{N}}$ et f dans $C^0([a,b],\mathbb{C})$. On dit que la suite (f_n) converge en moyenne vers f sur $[a,b]$ si et seulement si*

$$\lim_{n\to+\infty}\|f_n-f\|_1=\lim_{n\to+\infty}\int_a^b|f(t)-f_n(t)|\,dt=0$$

On dit que cette suite converge en moyenne quadratique sur $[a,b]$ si et seulement si

$$\lim_{n\to+\infty}\|f_n-f\|_2=\lim_{n\to+\infty}\left(\int_a^b|f(t)-f_n(t)|^2\,dt\right)^{\frac{1}{2}}=0$$

(ce qui équivaut évidemment à $\lim_{n\to+\infty}\|f_n-f\|_2^2=0$ et évite de traîner les racines carrées dans les calculs).

Il s'agit donc de la convergence respectivement dans les espaces vectoriels normés $(C^0([a,b],\mathbb{C}),\|\ \|_1)$ et $(C^0([a,b],\mathbb{C}),\|\ \|_2)$.

REMARQUE 10-1.19 On pourrait donner (et nous utiliserons ultérieurement, notamment dans le chapitre sur les séries de Fourier) une définition analogue en supposant simplement les fonctions continues par morceaux. Il y aurait alors un légère différence, due au fait que l'on travaille alors dans un espace **semi-normé** : si la suite (f_n) converge en moyenne (ou en moyenne quadratique) vers f, elle converge également vers toute fonction g se déduisant de f par modification en un nombre fini de points (et toute fonction g "limite" de la suite (f_n) est de ce type, puisque l'inégalité $\|f-g\|_1\leqslant\|f-f_n\|_1+\|g-f_n\|_1$ donne alors aisément $\|f-g\|_1=0$) : **il n'y a plus unicité de la limite**, ce qui n'est pas très gênant puisque f et g sont "indistinguables" vis à vis du calcul intégral.

PROPOSITION 10-1.20 Pour $f\in C_m^0([a,b],\mathbb{C})$, on a les inégalités :

$$\begin{cases}\|f\|_2\leqslant\sqrt{b-a}\,\|f\|_\infty\\ \|f\|_1\leqslant\sqrt{b-a}\,\|f\|_2\end{cases}$$

inégalités qui montrent que, dans $C^0([a,b],\mathbb{C})$ (et aussi dans $C_m^0([a,b],\mathbb{C})$, voir la remarque qui précède) la convergence uniforme entraîne la convergence en moyenne quadratique, qui elle-même entraîne la convergence en moyenne.

Démonstration : La première inégalité est conséquence de la majoration

$$\int_a^b|f(t)|^2\,dt\leqslant(b-a)\,\|f\|_\infty^2$$

la seconde découle de l'inégalité de Schwarz

$$\|f\|_1=\int_a^b|f(t)|.1\,dt\leqslant\left(\int_a^b|f(t)|^2\,dt\right)^{\frac{1}{2}}.\left(\int_a^b dt\right)^{\frac{1}{2}}\quad\blacksquare$$

Ces majorations montrent également que la convergence uniforme (*cf.* section **10-1.2.2**), la convergence en moyenne quadratique et la convergence en moyenne entraînent la convergence des intégrales, puisque par inégalité du module on a

$$\left|\int_a^b f(t)\, dt - \int_a^b f_n(t)\, dt\right| \leqslant \|f - f_n\|_1$$

REMARQUE 10-1.21 Par contre, on peut voir que la convergence en moyenne n'entraîne pas la convergence en moyenne quadratique : sur $\mathcal{C}^0([a,b],\mathbb{C})$, les normes $\|\ \|_1$ et $\|\ \|_2$ ne sont pas équivalente. Par exemple, la suite de fonctions

$$f_n(t) = \begin{cases} \sqrt{n}(1-nt) & \text{si} \quad 0 \leqslant t \leqslant \dfrac{1}{n} \\ 0 \quad \text{si} \quad \dfrac{1}{n} \leqslant t \leqslant 1 \end{cases}$$

converge vers 0 dans $(\mathcal{C}^0([0,1],\mathbb{C}), \|\ \|_1)$, mais pas dans l'espace $(\mathcal{C}^0([0,1],\mathbb{C}), \|\ \|_2)$ (exercice !)

REMARQUE 10-1.22 Comme la terminologie l'indique, les normes des convergences en moyenne tiennent compte des fonctions "globalement" sur $[a,b]$, et non de manière "ponctuelle". En particulier, la convergence en moyenne n'a pas vraiment de rapport avec la convergence simple[7] sur $[a,b]$: on peut construire sur $[0,1]$ une suite $(f_n)_{n\in\mathbb{N}}$ qui converge en moyenne, alors que, pour tout $t \in [0,1]$, la suite $(f_n(t))_{n\in\mathbb{N}}$ est divergente. En travaillant avec des fonctions continues par morceaux[8], on peut prendre la suite

$$f_1 = 1_{[0,1]},\ f_2 = 1_{[0,\frac{1}{2}]},\ f_3 = 1_{[\frac{1}{2},1]},\ f_4 = 1_{[0,\frac{1}{4}]},\ f_5 = 1_{[\frac{1}{4},\frac{1}{2}]}, \dots ,\ f_8 = 1_{[0,\frac{1}{8}]} \text{ etc ...}$$

obtenue en faisant glisser une marche d'escalier de longueur 2^{-n} le long de $[0,1]$. Il est clair que cette suite converge en moyenne vers la fonction nulle sur $[0,1]$ puisque

$$\forall n \geqslant 2^p \quad \|f_n\|_1 \leqslant \frac{1}{2^p}$$

Par contre, pour tout $t \in [0,1]$, la suite $(f_n(t))_{n\in\mathbb{N}}$ est divergente, puisqu'elle prend une infinité de fois les valeurs 0 et 1.

EXERCICE 10-1.23 On suppose traité l'exercice **10-1**.16. Pour p réel $\geqslant 1$ et pour $f \in \mathcal{C}_m^0([a,b],\mathbb{C})$, on pose

$$\|f\|_p = \left(\int_a^b |f(t)|^p\, dt\right)^{\frac{1}{p}}$$

Montrer que l'on définit ainsi une norme sur $\mathcal{C}^0([a,b],\mathbb{C})$ et une semi-norme sur l'espace $\mathcal{C}_m^0([a,b],\mathbb{C})$. La convergence associée à cette norme est dite "convergence en moyenne

[7]Cependant, une question qui nous occupera beaucoup ultérieurement sera la recherche d'hypothèses à rajouter à la convergence simple pour assurer la convergence en moyenne. La convergence uniforme est un exemple, qui a le défaut d'être techniquement difficile à manipuler.

[8]On pourrait modifier cet exemple par interpolation pour travailler avec des fonctions continues.

d'ordre p ". Montrer que, pour $1 \leqslant p_1 < p_2$, la convergence d'une suite en moyenne d'ordre p_2 entraîne la convergence de cette suite vers la même limite en moyenne d'ordre p_1. Montrer que tout fonction $f \in C_m^0([a,b], \mathbb{C})$ est limite en moyenne d'ordre p d'une suite de fonctions en escalier, d'une suite de fonctions continues, d'une suite de fonctions continues s'annulant en a et b.

EXERCICE 10-1.24 Si f est une fonction **continue** de $[a,b]$ dans $\mathbb{C}$, montrer que

$$\lim_{p \to +\infty} \|f\|_p = \|f\|_\infty$$

10-1.2.6 Sommes de Riemann d'une fonction continue

DÉFINITION 10-1.25 *Soit $f \in C^0([a,b], E)$. Pour une subdivision*

$$d : a = t_0 < t_1 < t_2 < \cdots < t_p = b$$

de $[a,b]$, on considère une famille de "points intermédiaires"

$$d' = (\zeta_i)_{0 \leqslant i \leqslant p-1} \quad \text{avec} \quad \forall i \quad t_i \leqslant \zeta_i \leqslant t_{i+1}$$

et on appelle somme de Riemann de la fonction f, pour la subdivision d et les points intermédiaires d', le vecteur de E

$$\Sigma(f, d, d') = \sum_{i=0}^{p-1} (t_{i+1} - t_i) f(\zeta_i)$$

C'est en fait l'intégrale d'une fonction en escalier définie sur $[a,b]$ qui prend la valeur $f(\zeta_i)$ sur $]t_i, t_{i+1}[$.

DÉFINITION 10-1.26 *On appelle pas de la subdivision*

$$d : a = t_0 < t_1 < t_2 < \cdots < t_p = b$$

la plus grande des longueurs des intervalles de cette subdivision :

$$\pi(d) = \max_{0 \leqslant i \leqslant p-1} (t_{i+1} - t_i)$$

THÉORÈME 10-1.27 Si $f : [a,b] \to E$ est continue, elle vérifie

$$\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall d \text{ subdivision de } [a,b] \quad \forall d' \text{ intermédiaire}$$
$$\pi(d) \leqslant \alpha \Rightarrow \left| \int_a^b f(t)\, dt - \Sigma(f, d, d') \right| \leqslant \varepsilon$$

Démonstration : f étant uniformément continue sur $[a,b]$, si $\varepsilon > 0$ est fixé, on peut trouver un $\alpha > 0$ tel que

$$\forall t, t' \in [a,b] \quad |t - t'| \leqslant \alpha \Rightarrow |f(t) - f(t')| \leqslant \frac{\varepsilon}{b-a}$$

Si $d : a = t_0 < t_1 < t_2 < \cdots < t_p = b$ est une subdivision de $[a,b]$ de pas inférieur à α, $d' = (\zeta_i)_{0 \leqslant i \leqslant p-1}$ des points intermédiaires, et s

une fonction en escalier définie sur $[a,b]$ qui prend la valeur $f(\zeta_i)$ sur $]t_i, t_{i+1}[$, on a

$$\left|\int_a^b f(t)\ dt - \Sigma(f,d,d')\right| = \left|\int_a^b f(t)\ dt - \int_a^b s(t)\ dt\right|$$
$$\leqslant \sum_{i=0}^{p-1} \int_{t_i}^{t_{i+1}} |f(t) - s(t)|\ dt$$

Comme sur $]t_i, t_{i+1}[$ on a

$$|f(t) - s(t)| = |f(t) - f(\zeta_i)| \leqslant \frac{\varepsilon}{b-a}$$

et puisque $|t - \zeta_i| \leqslant t_{i+1} - t_i \leqslant \pi(d) \leqslant \alpha$, on obtient aisément

$$\left|\int_a^b f(t)\ dt - \Sigma(f,d,d')\right| \leqslant \sum_{i=0}^{p-1} \frac{\varepsilon}{b-a}(t_{i+1} - t_i) = \varepsilon$$

ce qui démontre le résultat. On résume souvent ce dernier (de manière imprécise) en disant que "les sommes de Riemann d'une fonction continue sur un segment tendent vers l'intégrale de cette fonction lorsque le pas de subdivision tend vers 0 ". ■

En particulier, en prenant une subdivision de $[a,b]$ en n sous-intervalles de même longueur, on obtient facilement le

COROLLAIRE 10-1.28 Si $f : [a,b] \to E$ est continue, on a

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) = \frac{1}{b-a}\int_a^b f(t)\ dt$$

quantité que l'on appelle valeur moyenne de la fonction f sur le segment $[a,b]$.

EXERCICE 10-1.29 Démontrer que le théorème précédent et son corollaire subsistent pour une fonction f continue par morceaux sur $[a,b]$. (On pourra commencer par le vérifier pour la fonction caractéristique d'un sous-intervalle de $[a,b]$).

EXERCICE 10-1.30 Si $f \in C^0([0,1], E)$, déterminer

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} (-1)^k f\left(\frac{k}{n}\right)$$

EXERCICE 10-1.31 (INEGALITE DE JENSEN) Soit f une fonction continue par morceaux d'un intervalle $[a,b]$ à valeurs dans un intervalle I de $\mathbb{R}$. Montrer que

$$\frac{1}{b-a}\int_a^b f(t)\ dt \in I$$

Si $\varphi : I \to \mathbb{R}$ est une fonction convexe, montrer que

$$\varphi\left(\frac{1}{b-a}\int_a^b f(t)\ dt\right) \leqslant \frac{1}{b-a}\int_a^b \varphi(f(t))\ dt$$

Plus généralement, montrer que, si $\psi : [a,b] \to \mathbb{R}^+$ est une fonction continue par morceaux positive d'intégrale égale à 1

$$\varphi\left(\int_a^b f(t)\,\psi(t)\ dt\right) \leqslant \int_a^b \varphi(f(t))\,\psi(t)\ dt$$

Si $f > 0$ sur $[0,1]$, comparer

$$\ln\left(\int_0^1 f(t)\ dt\right) \quad \text{et} \quad \int_0^1 \ln f(t)\ dt$$

10-1.3 Intégrale fonction d'une de ses bornes

10-1.3.1 Relation de Chasles

DÉFINITION 10-1.32 *Si I est un intervalle non compact de $\mathbb{R}$, une fonction $f : I \to E$ est dite continue par morceaux ssi sa restriction à tout segment K inclus dans I est continue par morceaux.*

Sur tout segment inclus dans I, la fonction f ne présente qu'un nombre fini de points de discontinuité, qui sont tous de "première espèce".

DÉFINITION 10-1.33 *Si I est un intervalle quelconque et $f : I \to E$ est continue par morceaux, on définit*

$$\int_a^b f(t)\ dt = \begin{cases} \displaystyle\int_{[a,b]} f & si \quad a < b \\ 0_E & si \quad a = b \\ -\displaystyle\int_{[a,b]} f & si \quad a > b \end{cases}$$

On a donc toujours $\forall\, a,b \in I \quad \displaystyle\int_a^b f(t)\ dt = -\int_b^a f(t)\ dt$

Attention aux inégalités ! Si f est positive et $b < a$, $\displaystyle\int_a^b f(t)\ dt$ est négative.

PROPOSITION 10-1.34 Si $f : I \to E$ est continue par morceaux, on a

$$\forall\, a,b \textbf{ et } c \in I \qquad \int_a^b f(t)\ dt = \int_a^c f(t)\ dt + \int_c^b f(t)\ dt$$

Démonstration : La formule à démontrer s'écrit aussi

$$\int_a^c f(t)\ dt + \int_c^b f(t)\ dt + \int_b^a f(t)\ dt = 0_E$$

Elle est évidente si a, b et c ne sont pas tous distincts. Dans le cas général, le membre de gauche est évidemment invariant par permutation circulaire sur (a, b, c) et changé en son opposé par une transposition de deux de ces trois réels. La formule étant vérifiée lorsque $a \leqslant c \leqslant b$ (par additivité de l'intégrale par rapport à l'intervalle d'intégration), elle l'est également quelque soit l'ordre des trois réels a, b et c. ■

10-1.3.2 Continuité, dérivabilité

Si f est une fonction continue par morceaux $I \to E$ et a est un point fixé dans I, on étudie ici les propriétés de l'application

$$F : I \to E \quad \text{définie par} \quad x \mapsto F(x) = \int_a^x f(t)\, dt$$

THÉORÈME 10-1.35 Sous les hypothèses précédentes, F est localement lipschitzienne sur I (c'est-à-dire que sa restriction à tout segment inclus dans I est lipschitzienne). En particulier, F est continue sur I.

Démonstration : Soit K un segment inclus dans I. Comme $f|_K$ est continue par morceaux, elle est bornée sur le compact K. Il existe donc une constante $M_K > 0$ avec

$$\forall t \in K \quad \|f(t)\| \leqslant M_K$$

Pour $x, y \in K$, avec par exemple $x \leqslant y$, on a

$$\|F(y) - F(x)\| = \left\| \int_x^y f(t)\, dt \right\| \leqslant \int_x^y \|f(t)\|\, dt \leqslant M_K\, |y - x|$$

puisque $[x, y] \subset K$. ■

En fait, l'existence en tout point d'une limite à droite et à gauche pour f entraîne des propriétés de dérivabilité pour F :

THÉORÈME 10-1.36 F est continue et C^1 par morceaux sur I : en tout point x_0 intérieur à I la fonction F possède une dérivée à droite et à gauche, données par

$$F'_d(x_0) = \lim_{t \to x_0^+} f(t) = f(x_0 + 0) \quad \text{et} \quad F'_g(x_0) = \lim_{t \to x_0^-} f(t) = f(x_0 - 0)$$

De même, si I contient son extrémité gauche α, on aura $F'_d(\alpha) = f(\alpha + 0)$, avec un résultat analogue de dérivabilité à gauche en l'extrémité droite de I. En particulier, F est dérivable en tout point x de continuité de f, avec $F'(x) = f(x)$.

Démonstration : On se limitera à prouver la dérivabilité à droite en tout point dont I est voisinage à droite. Si $[x_0, x_0 + \eta[\subset I$, on a, pour $0 < h < \eta$

$$\delta(h) = F(x_0 + h) - F(x_0) - h\, f(x_0 + 0) = \int_{x_0}^{x_0 + h} [f(t) - f(x_0 + 0)]\, dt$$

Il s'agit de montrer que cette quantité est un $o(h)$ pour $h \to 0^+$. Ceci est conséquence de la définition de la limite à droite : si $\varepsilon > 0$ est fixé

$$\exists \beta > 0 \quad \forall t \in]x_0, x_0 + \beta] \quad \|f(t) - f(x_0 + 0)\| \leqslant \varepsilon$$

On a alors, par inégalité de la norme, pour $0 < h \leqslant \beta$

$$\|\delta(h)\| \leqslant \int_{x_0}^{x_0+h} \|f(t) - f(x_0 + 0)\| \, dt \leqslant \varepsilon h$$

ce qui prouve le résultat. ■

10-1.3.3 Primitive d'une fonction continue

DÉFINITION 10-1.37 *Soit I un intervalle de $\mathbb{R}$. Si $g : I \to E$ est une fonction quelconque, une fonction $G : I \to E$ est dite primitive de g sur I si et seulement si G est dérivable en tout point de I et si*

$$\forall x \in I \qquad G'(x) = g(x)$$

L'existence d'une telle fonction, pour g quelconque, n'est pas assurée[9]. Il est évident que, si g possède une primitive sur l'intervalle I, elle en possède une infinité, deux d'entre elles différant d'une fonction constante.

Lorsque la fonction est continue sur I, le calcul intégral assure l'existence de primitives : le théorème **10-1**.36 donne immédiatement

THÉORÈME 10-1.38 Si $f : I \to E$ est continue, avec I intervalle de $\mathbb{R}$, f admet sur I une infinité de primitives. Si $a \in I$ est fixé, une quelconque de ces primitives s'écrira

$$x \mapsto H(x) = \int_a^x f(t) \, dt + \lambda$$

où λ est un vecteur de E (évidemment égal à $H(a)$).

Nous obtenons (enfin !) un moyen "pratique" de calculer l'intégrale d'une fonction **continue sur un segment** : cette intégrale, définie au départ par un procédé de passage à la limite, pourra s'obtenir **par un calcul de primitive.**[10]

Plus précisément, on a le **théorème fondamental du calcul intégral** :

COROLLAIRE 10-1.39 Si $f : I \to E$ est continue, et F est une primitive de f sur I, on a

$$\forall a, b \in I \quad \int_a^b f(t) \, dt = F(b) - F(a) = [F(t)]_a^b$$

[9]Par exemple, si g est réelle, l'exercice **7-7**.14 montre qu'une condition nécessaire d'existence de G est que g vérifie la propriété des valeurs intermédiaires.

[10]Cette commodité ne doit pas faire oublier la vraie nature de l'intégrale (pour parler de manière imprécise : somme d'une infinité de termes infiniment petits).

Démonstration : évident, puisque la fonction $x \mapsto F(x) - \int_a^x f(t)\, dt$ est constante sur I. ■

Si f est seulement supposée continue par morceaux, on utilisera la relation de Chasles en travaillant avec une subdivision de $[a,b]$ (ou $[b,a]$) adaptée à $f|_{[a,b]}$: si x_i et x_{i+1} sont deux points de subdivision consécutifs, on évaluera $\int_{x_i}^{x_{i+1}} f(t)\, dt$, en travaillant avec une primitive de la fonction continue $f|_{]x_i,x_{i+1}[}$ (prolongeable en une fonction continue sur $[x_i, x_{i+1}]$).

REMARQUE 10-1.40 La formule précédente peut s'écrire aussi

$$\int_a^b g'(t)\, dt = g(b) - g(a)$$

Elle est valable pour g de classe C^1 sur $[a,b]$ (car g est alors primitive de la fonction continue g'). Elle est valable avec un petit abus de notation dans le cas où g est continue et C^1 par morceaux sur $[a,b]$ (voir à cet effet le commentaire qui débute la démonstration du corollaire **10-1**.45).

REMARQUE 10-1.41 Si f est une fonction continue par morceaux sur I, on dit parfois (avec un petit abus de langage) que

$$F(x) = \int_a^x f(t)\, dt$$

est une primitive de f sur I. Cela signifie que F est une fonction continue et C^1 par morceaux sur I, dérivable en chaque point x de continuité de f, avec $F'(x) = f(x)$.

10-1.4 Calculs d'intégrales

Pour les fonctions continues, le calcul d'intégrales revient donc à "inverser l'opération de dérivation" :

10-1.4.1 Intégration par parties

Nous énonçons un résultat général, sachant bien que nous l'appliquerons le plus souvent avec des fonctions complexes et pour l'opération de produit usuel sur de telles fonctions.

THÉORÈME 10-1.42 Soient E, F, G trois espaces normés complets, et $B : E \times F \to G$ une application bilinéaire continue. Si

$$u : [a,b] \to E \quad \textbf{et} \quad v : [a,b] \to F$$

sont deux applications de classe C^1, on a

$$\int_a^b B(u(t), v'(t))\, dt = [B(u(t), v(t))]_a^b - \int_a^b B(u'(t), v(t))\, dt$$

Démonstration : La fonction $t \mapsto B(u(t), v'(t)) + B(u'(t), v(t))$ est continue sur $[a, b]$, et possède la fonction $t \mapsto B(u(t), v(t))$ comme primitive. On lui applique la formule fondamentale du calcul intégral. ■

Dans le cas du produit usuel de fonctions complexes, cette formule s'écrira

$$\int_a^b u(t) v'(t)\, dt = [u(t) v(t)]_a^b - \int_a^b u'(t) v(t)\, dt$$

Cette formule est d'un usage très fréquent : elle permet de transformer une intégrale en une autre, supposée "plus simple". Plus simple souvent pour un calcul explicite, mais souvent aussi plus simple à majorer, à évaluer en terme de comportement asymptotique lorsque l'on fait varier b etc... L'intégration par parties est un outil qu'il faut avoir à l'esprit, tout l'art étant de découvrir les bonnes fonctions u et v qui vont effectivement faire retomber sur quelque chose de "plus simple".

EXERCICE 10-1.43 (INTEGRALES DE WALLIS) Pour n entier, on pose

$$I_n = \int_0^{\frac{\pi}{2}} \sin^n t\, dt$$

Montrer que, pour $n \geqslant 2$

$$I_n = \frac{(n-1)}{n} I_{n-2}$$

et en déduire les valeurs de I_{2p} et I_{2p+1} (qu'on s'efforcera d'écrire en utilisant des factorielles). Montrer que la suite (I_n) est décroissante et tend vers 0, et que

$$\lim_{n \to +\infty} \frac{I_{n+1}}{I_n} = 1$$

En déduire que la valeur de la constante e^L intervenant dans la formule de Stirling (*cf.* section **9-2.4.3**) est effectivement $\sqrt{2\pi}$.

EXERCICE 10-1.44 Trouver un équivalent, pour $x \to +\infty$, de

$$F(x) = \int_e^x \ln(\ln(t))\, dt$$

On peut affaiblir légèrement les hypothèses de la formule d'intégration par parties. En tout état de cause, une intégration par parties doit toujours être justifiée avec concision et précision. Le corollaire qui suit nous sera en particulier très utile dans le chapitre consacré aux séries de Fourier : il s'agit de la généralisation aux fonctions **continues et C^1 par morceaux** (*cf.* définition **8-4**.23) :

COROLLAIRE 10-1.45 La formule d'intégration par parties reste valable lorsque u et v sont des fonctions <u>continues</u> et C^1 par morceaux.

Démonstration : dans la formule d'intégration par parties, il y a alors un petit abus d'écriture : les fonctions u' et v' sont définies sur $[a,b]$, sauf en un nombre fini de points. Mais possédant en ces points des limites à droite et à gauche, elles sont prolongeables en des fonctions continues par morceaux sur $[a,b]$, ce qui permet de donner un sens aux intégrales $\int_{[a,b]} B(u,v')$ et $\int_{[a,b]} B(u',v)$, dont les valeurs ne dépendent pas des prolongements effectivement choisis pour u' et v'. Pour la démonstration, il suffit de prendre une subdivision de $[a,b]$

$$d : a = t_0 < t_1 < t_2 < \cdots < t_p = b$$

adapté à la fois à u et v. On écrit alors

$$\int_a^b B(u(t), v'(t))\, dt = \sum_{i=0}^{p-1} \int_{t_i}^{t_{i+1}} B(u(t), v'(t))\, dt$$

chacune de ces intégrales pouvant être évaluée par parties, en travaillant avec les prolongements C^1 de u et v à $[t_i, t_{i+1}]$:

$$\int_{t_i}^{t_{i+1}} B(u(t), v'(t))\, dt = B\left(u\left(t_{i+1}^-\right), v\left(t_{i+1}^-\right)\right) - B\left(u\left(t_i^+\right), v\left(t_i^+\right)\right) - \int_{t_i}^{t_{i+1}} B(u'(t), v(t))\, dt$$

formule valable si u et v sont simplement C^1 par morceaux. Comme on rajoute l'hypothèse de continuité, la partie "toute intégrée" s'écrit en fait

$$B(u(t_{i+1}), v(t_{i+1})) - B(u(t_i), v(t_i))$$

et donne une somme télescopique. On obtient finalement

$$\int_a^b B(u(t), v'(t))\, dt = [B(u,v)]_a^b - \sum_{i+0}^{p-1} \int_{t_i}^{t_{i+1}} B(u'(t), v(t))\, dt = [B(u,v)]_a^b - \int_a^b B(u'(t), v(t))\, dt$$

On remarque qu'il faudrait tenir compte des sauts de u et v si on ne faisait pas d'hypothèse de continuité. ■

10-1.4.2 Formule de Taylor avec reste sous forme d'intégrale

La formule d'intégration par parties permet d'obtenir une écriture du terme complémentaire de la formule de Taylor-Lagrange, que nous avons déjà utilisée dans une démonstration "simplifiée" de cette formule (*cf.* section **8-5.3.3**) :

THÉORÈME 10-1.46 Si $n \in \mathbb{N}$ et $f : [a,b] \to E$ est de classe C^{n+1}, on a

$$f(b) = f(a) + \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\, dt$$

Cette formule est encore valable si $b < a$, ou si on suppose simplement f de classe C^n et C^{n+1} par morceaux sur $[a,b]$.

Démonstration : La démonstration se fait par récurrence sur n. Pour $n = 0$, c'est simplement la formule de la remarque **10-1**.40

$$\int_a^b f'(t)\, dt = f(b) - f(a)$$

valable pour f continue et C^1 par morceaux. Nous supposons la formule vérifiée lorsque f est de classe C^n et C^{n+1} par morceaux sur $[a,b]$, et prenons à présent f de classe C^{n+1} et C^{n+2} par morceaux sur $[a,b]$. Par hypothèse de récurrence, on a

$$f(b) = f(a) + \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\, dt$$

et on transforme la dernière intégrale par parties

$$\int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\, dt = \left[-\frac{(b-t)^{n+1}}{(n+1)!} \right]_a^b + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\, dt$$

(car $f^{(n+1)}$ est continue et C^1 par morceaux sur $[a,b]$). On obtient exactement la formule à démontrer, au rang $n+1$. ■

REMARQUE 10-1.47 L'expérience montre que l'on a souvent du mal à mémoriser l'expression du terme complémentaire de la formule de Taylor sous forme d'intégrale. Pour éviter les erreurs, se souvenir que le polynôme qui intervient sous l'intégrale est en $(b-t)$, avec b "borne supérieure" de l'intégrale, et tester la formule à l'ordre n lorsque f est un polynôme de degré $n+1$, c'est-à-dire lorsque la fonction $f^{(n+1)}$ est constante égale à λ. On sait dans ce cas que la formule de Taylor pour les polynômes donne le terme complémentaire

$$\lambda \frac{(b-a)^{n+1}}{(n+1)!}$$

La fonction sous le signe intégrale ne peut donc être que

$$\frac{(b-t)^{\boxed{n}}}{\boxed{n}!} f^{\boxed{(n+1)}}(t)$$

REMARQUE 10-1.48 L'inégalité de Taylor-Lagrange

$$\left\| f(b) - f(a) - \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) \right\| \leqslant \frac{|b-a|^{n+1}}{(n+1)!} \sup_{[a,b]} \left\| f^{(n+1)} \right\|$$

compare la norme de la différence

$$f(b) - f(a) - \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a)$$

à $|b-a|^{n+1}$. Il est donc parfois intéressant de faire apparaître $(b-a)^{n+1}$ en facteur dans le reste intégral. Ceci se fait par le changement de variable (en anticipant les résultats bien connus de la section qui suit)

$$[0,1] \ni u \mapsto t = a + u(b-a)$$

On obtient alors facilement, pour f de classe C^{n+1}

$$f(b) = f(a) + \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + (b-a)^{n+1} \int_0^1 \frac{(1-u)^n}{n!} f^{(n+1)}(a + u(b-a))\, du$$

10-1.4.3 Intégration par "changement de variable"

THÉORÈME 10-1.49 Soit I un intervalle de $\mathbb{R}$, et f continue : $I \to E$. Soit φ une application de classe C^1 définie sur un segment $[\alpha, \beta]$ à valeurs dans I. On a alors

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(u)\, du = \int_{\alpha}^{\beta} f(\varphi(t))\, \varphi'(t)\, dt \qquad (*)$$

Démonstration : Les deux membres de cette égalité ont un sens : $[\varphi(\alpha), \varphi(\beta)]$ (ou $[\varphi(\beta), \varphi(\alpha)]$) est un intervalle sur lequel f est continue. De même, comme $\varphi([\alpha, \beta]) \subset I$, l'application

$$\psi : t \mapsto f(\varphi(t))\, \varphi'(t)$$

est continue sur $[\alpha, \beta]$. Si F est une primitive de f sur I (une telle fonction existe, puisqu'on a supposé f continue), la formule de dérivation d'une fonction composée montre que $F \circ \varphi$ est primitive sur $[\alpha, \beta]$ de la fonction continue ψ. On a donc

$$\int_{\alpha}^{\beta} \psi(t)\, dt = F \circ \varphi(\beta) - F \circ \varphi(\alpha) = [F]_{\varphi(\alpha)}^{\varphi(\beta)} = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(u)\, du \quad \blacksquare$$

Quelques remarques s'imposent au sujet de cette formule :

- Dans l'égalité (*), le membre de gauche a un sens dès que f est continue sur $[\varphi(\alpha), \varphi(\beta)]$. Le membre de droite en a un si f est continue sur $\varphi([\alpha, \beta])$. Ce n'est pas la même chose[11] ! Pour transformer "mécaniquement" l'intégrale

$$\int_{a=\varphi(\alpha)}^{b=\varphi(\beta)} f(u)\, du \quad \text{en} \quad \int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t)\, dt$$

avec $a = \varphi(\alpha)$ et $b = \varphi(\beta)$ et φ de classe $\mathcal{C}^1$ sur $[\alpha, \beta]$, en "posant $u = \varphi(t)$" comme on le dit souvent (et c'est là l'intérêt des notations : on remplace du par $\varphi'(t)\, dt$), il ne faut pas l'oublier.

- Il n'est pas utile de supposer φ injective sur $[\alpha, \beta]$: cela n'intervient pas dans la démonstration précédente. La terminologie "intégration par changement de variable" est assez malheureuse et est à l'origine de confusions. Un "bon" changement de variable, utilisé pour passer d'un problème à un autre, doit souvent permettre de faire l'aller et le retour, donc doit être bijectif. Ce n'est pas nécessaire pour appliquer la formule précédente. Par exemple, pour $f : [0, 1] \to \mathbb{C}$ continue, on a

$$\int_0^1 f(u)\, du = \int_0^{\frac{\pi}{2}} f(\sin t) \cos t\, dt$$

Si f est de plus continue sur $[-1, 1]$, on aura également

$$\int_0^1 f(u)\, du = \int_0^{\frac{\pi}{2}} f(\sin t) \cos t\, dt = \int_0^{\frac{5\pi}{2}} f(\sin t) \cos t\, dt$$

- En réfléchissant un peu, on s'aperçoit qu'on peut remplacer l'hypothèse "φ de classe $\mathcal{C}^1$" par "φ est continue et $\mathcal{C}^1$ par morceaux". Par contre, on ne peut sans précaution remplacer "f continue" par "f continue par morceaux". En effet, sans hypothèse supplémentaire sur φ, on n'est alors plus assuré du fait que $(f \circ \varphi)\varphi'$ soit continue par morceaux. Prendre par exemple sur $I = [-1, 1]$ la fonction f égale à la fonction caractéristique de $[0, 1]$ (c'est une bonne fonction en escalier) et $\varphi : [-1, 1] \to [-1, 1]$ définie par

$$\varphi(t) = t^4 \sin\left(\frac{1}{t}\right) \quad \text{avec} \quad \varphi(0) = 0$$

On vérifie aisément que c'est une fonction de classe $\mathcal{C}^1$. La fonction $(f \circ \varphi)\varphi'$ n'est pas continue par morceaux sur $[-1, 1]$, puisqu'elle vaut (en dehors de l'origine)

$$\varphi'(t) = 4t^3 \sin\left(\frac{1}{t}\right) - t^2 \cos\left(\frac{1}{t}\right)$$

en tout point où $\sin\left(\frac{1}{t}\right) \geqslant 0$, et vaut 0 ailleurs. Elle présente une infinité de discontinuités sur $[-1, 1]$ (notamment tous les points de la forme $\frac{1}{2k\pi}$,

[11] C'est évidemment la même chose si φ est $\mathcal{C}^1$ et monotone !

avec $k \in \mathbb{Z}$). Avec la présentation de l'intégrale qui est la nôtre, on ne peut donner un sens à

$$\int_{-1}^{1} f(\varphi(t))\,\varphi'(t)\,dt$$

Par contre, l'intégrale

$$\int_{\varphi(-1)}^{\varphi(1)} f(u)\,du = \int_{-\sin 1}^{\sin 1} f(u)\,du = \sin 1$$

existe.

Si l'on veut une formule de changement de variable avec f **continue par morceaux**, on supposera que φ **est un "bon changement de variable", c'est-à-dire est strictement monotone.** Ce sera en particulier le cas pour des opérations simples comme les translations ($\varphi(t) = t - t_0$), les symétries ($\varphi(t) = t_0 - t$) et les homothéties ($\varphi(t) = \lambda(t - t_0)$) :

THÉORÈME 10-1.50 Si f est continue par morceaux de I dans E, et φ est une fonction de classe C^1 de $[\alpha,\beta] \to I$ strictement monotone, alors

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(u)\,du = \int_{\alpha}^{\beta} f(\varphi(t))\,\varphi'(t)\,dt$$

Démonstration : Faisons la démonstration dans le cas où φ est strictement décroissante. Montrons d'abord que $t \mapsto f(\varphi(t))\,\varphi'(t)$ est continue par morceaux sur $[\alpha,\beta]$. La fonction $\varphi|_{[\alpha,\beta]}$ est bijective de $[\alpha,\beta]$ dans $\varphi([\alpha,\beta]) = [\varphi(\beta), \varphi(\alpha)]$. Comme f est continue par morceaux, on peut choisir une subdivision

$$d : u_0 = \varphi(\beta) < u_1 < \cdots < u_p = \varphi(\alpha)$$

adaptée à $f|_{[\varphi(\beta),\varphi(\alpha)]}$. On lui fait correspondre (par l'intermédiaire de φ^{-1}) une subdivision de $[\alpha,\beta]$:

$$d' : t_0 = \alpha < t_1 = \varphi^{-1}(u_{p-1}) < \cdots < t_p = \beta$$

Posons $g = f \circ \varphi$. Sur $]t_i, t_{i+1}[$, la fonction g est continue (comme composée de $\varphi|_{]t_i,t_{i+1}[} :]t_i, t_{i+1}[\to]u_{p-i-1}, u_{p-i}[$ et de la fonction continue $f|_{]u_{p-i-1},u_{p-i}[}$. Elle est prolongeable en une fonction continue sur $[t_i, t_{i+1}]$, puisqu'on a clairement

$$g(t_i + 0) = f(u_{p-i} - 0) \quad \text{et} \quad g(t_{i+1} - 0) = f(u_{p-i-1} + 0)$$

La fonction g est donc bien continue par morceaux sur $[\alpha,\beta]$, et il en est de même de son produit par la fonction continue φ'. Pour évaluer l'intégrale de $t \mapsto f(\varphi(t))\,\varphi'(t)$ sur $[\alpha,\beta]$, on utilisera la relation de Chasles :

$$\int_{\alpha}^{\beta} f(\varphi(t))\,\varphi'(t)\,dt = \sum_{i=0}^{p-1} \int_{t_i}^{t_{i+1}} f(\varphi(t))\,\varphi'(t)\,dt$$

et on peut transformer chacune de ces intégrales par changement de variable, en considérant le prolongement continu de $f|_{]u_{p-i-1},u_{p-i}[}$ au segment $[u_{p-i-1}, u_{p-i}] = \varphi([t_i, t_{i+1}])$. On a alors, en notant encore f ce prolongement,

$$\int_{t_i}^{t_{i+1}} f(\varphi(t))\varphi'(t)\,dt = \int_{\varphi(t_i)}^{\varphi(t_{i+1})} f(u)\,du = \int_{u_{p-i}}^{u_{p-i-1}} f(u)\,du$$

ce qui donnera bien en sommant

$$\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt = \int_{u_p}^{u_0} f(u)\,du = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(u)\,du \quad \blacksquare$$

REMARQUE 10-1.51 La dérivée φ' peut s'annuler en certains points de $[\alpha,\beta]$. Il n'est pas indispensable que φ soit un C^1 difféomorphisme de $[\alpha,\beta]$ sur son image.

REMARQUE 10-1.52 Comme φ est cette fois un véritable changement de variable, on peut écrire, pour a et $b \in \varphi([\alpha,\beta])$ quelconques

$$\int_a^b f(u)\,du = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)\,dt$$

EXEMPLE 10-1.53 Si f est une fonction continue par morceaux sur un intervalle $[a,b]$, on a

$$\forall T \in \mathbb{R} \quad \int_a^b f(t)\,dt = \int_{a+T}^{b+T} f(t-T)\,dt$$

EXERCICE 10-1.54 Si $f : \mathbb{R} \to E$ est continue par morceaux et T-périodique, montrer que

$$\int_a^{a+T} f(u)\,du$$

ne dépend pas de a, et montrer que

$$\frac{1}{T}\int_0^T f(u)\,du = \lim_{x\to+\infty} \frac{1}{x}\int_0^x f(u)\,du$$

A quelle condition la fonction $x \mapsto \displaystyle\int_0^x f(u)\,du$ est-elle T-périodique ?

10-1.4.4 Théorème de relèvement

Nous nous intéressons ici aux intégrales de la forme

$$J = \int_a^b \frac{f'(t)}{f(t)}dt$$

où la fonction $f : [a,b] \to \mathbb{C}$ est de classe C^1 **et ne s'annule pas**. Ces hypothèses assurent la continuité de la fonction à intégrer et l'existence de cette intégrale.

Si f est à valeurs réelles, elle garde un signe constant sur $[a,b]$, et on sait qu'alors la fonction $F(t) = \ln(|f(t)|)$ est une primitive de la fonction $\dfrac{f'}{f}$ et par conséquent

$$J = \int_a^b \frac{f'(t)}{f(t)}dt = \ln(|f(b)|) - \ln(|f(a)|) = \ln\left(\frac{f(b)}{f(a)}\right)$$

Lorsque f est complexe, que se passe-t-il ? Comme $f(t) = x(t) + i\,y(t)$ (avec x et y de classe $\mathcal{C}^1$ à valeurs réelles) ne s'annule pas, la fonction

$$\rho(t) = |f(t)| = \sqrt{x^2(t) + y^2(t)}$$

est de classe $\mathcal{C}^1$ sur $[a,b]$.

Admettons qu'il existe une détermination de classe $\mathcal{C}^1$

$$[a,b] \ni t \mapsto \theta(t) \in \mathbb{R}$$

de l'argument de $f(t)$. Un calcul élémentaire donne alors, à partir de l'égalité $f(t) = \rho(t)\,e^{i\theta(t)}$

$$\frac{f'}{f} = \frac{\rho'}{\rho} + i\,\theta'$$

et on obtient, par linéarité de l'intégrale,

$$J = \int_a^b \frac{f'(t)}{f(t)}dt = \ln\frac{\rho(b)}{\rho(a)} + i\,(\theta(b) - \theta(a))$$

Cette expression est évidemment à rapprocher de l'expression d'une détermination du logarithme d'un nombre complexe non nul (*cf.* section **9-6.2.4**).

EXERCICE 10-1.55 Si $z = \alpha + i\,\beta$ est un nombre complexe de partie imaginaire β non nulle, déterminer

$$\lim_{x \to +\infty} \int_{-x}^{x} \frac{dt}{t - z}$$

Nous avons supposé l'existence d'une détermination $\mathcal{C}^1$ de l'argument de $f(t)$. Si celle-ci est intuitivement évidente (l'existence locale ne pose pas vraiment de problème, l'existence globale est moins claire), il faut une démonstration. Il est remarquable que la calcul précédent, pris à l'envers, nous donne la solution du problème. Commençons par étudier le cas d'une fonction de module 1, le cas général s'en déduisant simplement :

THÉORÈME 10-1.56 (RELEVEMENT D'UNE FONCTION $\mathcal{C}^1$) **Soit I un intervalle de $\mathbb{R}$ et une application de classe $\mathcal{C}^1$**

$$I \ni t \mapsto z(t) \in U$$

à valeurs dans l'ensemble des nombres complexes de module 1. **Il existe une application de classe $\mathcal{C}^1$**

$$I \ni t \mapsto \theta(t) \in \mathbb{R} \quad \textbf{avec} \quad \forall t \in I \quad z(t) = e^{i\theta(t)}$$

Si $t_0 \in I$ et $\theta_0 \in \mathbb{R}$ est une détermination de l'argument de $z(t_0)$, on peut prendre[12]

$$\forall t \in I \quad \theta(t) = \theta_0 - i \int_{t_0}^{t} \frac{z'(u)}{z(u)}\, du = \theta_0 + \int_{t_0}^{t} \operatorname{Im} \frac{z'(u)}{z(u)}\, du$$

Si z est de classe $\mathcal{C}^k$ avec $k \geqslant 2$, il en est évidemment de même de θ.

Démonstration : En dérivant la relation

$$\forall t \in I \quad |z(t)|^2 = z(t)\overline{z(t)} = 1$$

on obtient

$$\forall t \in I \quad z'(t)\overline{z(t)} + z(t)\overline{z'(t)} = 0 = 2\operatorname{Re}\left(z'(t)\overline{z(t)}\right) = 2\operatorname{Re}\frac{z'(t)}{z(t)}$$

ce qui montre que le rapport $\dfrac{z'}{z}$ est imaginaire pur, et donc

$$\operatorname{Im}\frac{z'(u)}{z(u)} = -i\frac{z'(u)}{z(u)}$$

(interprétation géométrique : la tangente au cercle est orthogonale au rayon). Montrons que la fonction θ définie plus haut répond à la question : elle est évidemment de classe $\mathcal{C}^1$, comme intégrale d'une fonction continue dépendant d'une de ses bornes. Il reste à montrer que

$$t \mapsto z(t)\, e^{-i\theta(t)}$$

est constante égale à 1. Comme par construction elle vaut 1 en t_0, il suffit de démontrer que sa dérivée est identiquement nulle sur I.

$$\forall t \in I \quad \left(z(t)\, e^{-i\theta(t)}\right) = e^{-i\theta(t)}\left[z'(t) - i\,\theta'(t)\, z(t)\right] = 0$$

puisque $\theta'(t) = -i\dfrac{z'(t)}{z(t)}$. ■

Nous utiliserons en particulier ce résultat dans les études métriques des arcs réguliers de classe $\mathcal{C}^k$ (avec $k \geqslant 2$) du plan euclidien. Pour définir la courbure, il faudra faire intervenir une fonction de classe $\mathcal{C}^{k-1}$ représentant une détermination de l'angle polaire de la tangente orientée. C'est le théorème de relèvement qui nous assurera l'existence d'une telle fonction.

[12] Si $t \mapsto \theta_1(t)$ est une autre détermination de cet argument, on doit avoir

$$\forall t \in I \quad \theta(t) - \theta_1(t) \in 2\pi\mathbb{Z}$$

et, par un argument de connexité,

$$\exists k \in \mathbb{Z} \quad \forall t \in I \quad \theta_1(t) = \theta(t) + 2k\pi$$

COROLLAIRE 10-1.57 Si z est une fonction de classe C^k ($k \geq 1$) d'un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{C}^*$, il existe une détermination C^k de $t \mapsto \arg(z(t))$. Si θ_0 est une détermination de $\arg(z(t_0))$, on peut prendre

$$\theta(t) = \theta_0 + \int_{t_0}^{t} \operatorname{Im} \frac{z'(u)}{z(u)}\, du$$

Démonstration : Il suffit de travailler avec

$$Z(t) = \frac{z(t)}{|z(t)|} = \frac{z(t)}{\rho(t)}$$

et de remarquer que

$$\frac{Z'}{Z} = \frac{z'}{z} - \frac{\rho'}{\rho}$$

a même partie imaginaire que $\dfrac{z'}{z}$. ■

10-2 Intégrales dépendant d'un paramètre

10-2.1 Intégrales à paramètres

Dans toute cette section, on considèrera une partie X d'un espace normé $(Y, \| \|)$. X est ensemble des valeurs d'un paramètre x. Dans la pratique, on aura souvent $Y = \mathbb{R}$ (ou parfois $\mathbb{R}^p$), et X sera en général un intervalle de $\mathbb{R}$. On considère également un segment $[a, b] \subset \mathbb{R}$ et une application

$$f : X \times [a, b] \to E \qquad (x, t) \mapsto f(x, t)$$

telle que, pour tout $x \in X$, l'application partielle $f(x, \bullet)$ soit continue (exceptionnellement continue par morceaux) sur $[a, b]$. On pourra donc définir son intégrale sur ce segment, et ainsi définir une fonction $F : X \to E$ par

$$\forall x \in X \quad F(x) = \int_a^b f(x, t)\, dt$$

(bien noter que t est ici une variable muette).

Nous nous intéressons ici aux propriétés de continuité et, lorsque X est un intervalle de $\mathbb{R}$, de dérivabilité de la fonction F. Les théorèmes qui suivent donneront des conditions suffisantes portant sur f qui assureront cette continuité ou cette dérivabilité. Dans un chapitre ultérieur sur l'intégration sur un intervalle non compact, nous verrons d'autres théorèmes, de portée plus générale.

10-2.2 Continuité

THÉORÈME 10-2.1 Si X est une partie d'un espace normé Y et si f est une application continue $X \times [a, b] \to E$, alors

$$F(x) = \int_a^b f(x, t)\, dt$$

définit une fonction continue sur X.

Démonstration : On prouve la continuité de F en un point $y \in X$ quelconque à l'aide du critère séquentiel : on montre que, pour toute suite $(x_n)_{n\in\mathbb{N}}$ de points de X,

$$\lim_{n\to+\infty} x_n = y \Rightarrow \lim_{n\to+\infty} F(x_n) = F(y)$$

On évalue

$$\begin{aligned} \|F(y) - F(x_n)\| \\ &= \left\| \int_a^b f(y,t)\, dt - \int_a^b f(x_n,t)\, dt \right\| \\ &\leqslant \int_a^b \|f(y,t) - f(x_n,t)\|\, dt \end{aligned}$$

La suite $(x_n)_{n\in\mathbb{N}}$ étant convergente vers y, on sait que

$$K = \{x_n,\, n \in \mathbb{N}\} \cup \{y\}$$

est un compact de Y. On en déduit que $K \times [a,b]$ est un compact de $Y \times \mathbb{R}$, sur lequel la fonction f est uniformément continue : si $\varepsilon > 0$ est donné, on peut trouver un $\alpha > 0$ tel que

$$\begin{aligned} &\forall z_1, z_2 \in K \quad \forall s,t \in [a,b] \\ &\qquad (\|z_1 - z_2\| \leqslant \alpha \text{ et } |s-t| \leqslant \alpha) \Rightarrow \|f(z_1,s) - f(z_2,t)\| \leqslant \varepsilon \end{aligned}$$

Pour cette valeur de α, puisque la suite x_n converge vers y, on peut trouver un rang N_α tel que

$$\forall n \geqslant N_\alpha \quad \|y - x_n\| \leqslant \alpha$$

Pour $n \geqslant N_\alpha$, on aura alors

$$\forall t \in [a,b] \quad \|f(x_n,t) - f(y,t)\| \leqslant \varepsilon$$

ce qui donne

$$\forall n \geqslant N_\alpha \quad \|F(y) - F(x_n)\| \leqslant \varepsilon(b-a)$$

et prouve bien la continuité de F. ■

10-2.3 Dérivabilité

On suppose ici que X est un intervalle de $\mathbb{R}$, et $f : X \times [a,b] \to E$ est une fonction continue, ce qui assure déjà la continuité de F sur X.

RAPPEL : Soit $x_0 \in X$ et $t_0 \in [a,b]$. On dit que la fonction $(x,t) \to f(x,t)$ possède, au point (x_0,t_0), une dérivée partielle par rapport à la première variable si et seulement si l'application partielle

$$X \to E \qquad x \mapsto f(x,t_0)$$

est dérivable en x_0. Sa dérivée en ce point sera notée

$$\frac{\partial f}{\partial x}(x_0,t_0) = \lim_{h\to 0} \frac{f(x_0+h,t_0) - f(x_0,t_0)}{h} \in E$$

THÉORÈME 10-2.2 Soit f continue de $X \times [a,b] \to E$, admettant en tout point de $X \times [a,b]$ une dérivée partielle par rapport à la première variable. On suppose de plus que la fonction "dérivée partielle"

$$\frac{\partial f}{\partial x} : X \times [a,b] \to E \qquad (x,t) \mapsto \frac{\partial f}{\partial x}(x,t)$$

est continue[13] sur $X \times [a,b]$. La fonction

$$F(x) = \int_a^b f(x,t)\, dt$$

est alors de classe C^1 sur X, et on a

$$\forall x \in X \qquad F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)\, dt$$

On dit souvent que cette dérivée est obtenue par dérivation sous le signe d'intégration.

Démonstration : Notons d'abord que les hypothèses faites sur f assurent l'existence de l'intégrale

$$\int_a^b \frac{\partial f}{\partial x}(x,t)\, dt$$

pour tout x de X, puisqu'il s'agit là de l'intégrale d'une fonction continue sur un segment. Montrons que F est dérivable en un point $x_0 \in X$ quelconque, et que sa dérivée peut être obtenue par dérivation sous le signe somme en prouvant que, pour $h \to 0$

$$\Delta(h) = F(x_0 + h) - F(x_0) - h. \int_a^b \frac{\partial f}{\partial x}(x_0,t)\, dt = o(h)$$

On a, par linéarité de l'intégrale,

$$\Delta(h) = \int_a^b \left[f(x_0 + h, t) - f(x_0,t) - h \frac{\partial f}{\partial x}(x_0,t) \right] dt$$

Par définition de la dérivée (partielle), à $t \in [a,b]$ fixé, la quantité entre [] est un $o(h)$ pour $h \to 0$. Ceci est insuffisant pour conclure immédiatement que $\Delta(h) = o(h)$. En effet, si $\varepsilon > 0$ est fixé, on peut, pour t fixé, trouver un $\alpha_t > 0$ tel que

$$|h| \leqslant \alpha_t \Rightarrow \left\| f(x_0 + h, t) - f(x_0,t) - h \frac{\partial f}{\partial x}(x_0,t) \right\| \leqslant \varepsilon h$$

[13] Il y a ici un abus d'écriture sur lequel nous reviendrons dans le chapitre de calcul différentiel en dimension finie : le symbole $\frac{\partial f}{\partial x}$ représente une fonction, définie par

$$\forall x_0 \in X \quad \forall t_0 \in [a,b] \quad \frac{\partial f}{\partial x}(x_0,t_0) = \text{dérivée en } x_0 \text{ de } f(\bullet,t_0)$$

Quand on écrit (x,t) le point générique de $X \times [a,b]$, ce "x" n'a rien à voir avec le symbole x apparaissant dans $\frac{\partial f}{\partial x}$.

Si l'on veut intégrer ensuite cette inégalité sur $[a, b]$, il faudrait trouver un α_t indépendant de t. Ceci peut-être fait avec un raisonnement plus précis :

Supposons par exemple x_0 intérieur à X, et considérons $\eta_0 > 0$ avec

$$[x_0 - \eta_0, x_0 + \eta_0] \subset X$$

Pour t fixé dans $[a, b]$, considérons la fonction

$$\varphi_t : [-\eta_0, \eta_0] \to E \qquad s \mapsto f(x_0 + s, t) - f(x_0, t) - s\frac{\partial f}{\partial x}(x_0, t)$$

Cette fonction est dérivable sur $[-\eta_0, \eta_0]$, puisqu'à une translation près sur la variable, il s'agit d'une application partielle de f à laquelle on a retranché une application affine. Par définition de la dérivée partielle par rapport à la première variable, on a

$$\forall s \in [-\eta_0, \eta_0] \quad \varphi_t'(s) = \frac{\partial f}{\partial x}(x_0 + s, t) - \frac{\partial f}{\partial x}(x_0, t)$$

Sur le compact $[x_0 - \eta_0, x_0 + \eta_0] \times [a, b]$, la fonction continue $\dfrac{\partial f}{\partial x}$ est uniformément continue. Si $\varepsilon > 0$ est choisi arbitrairement, on peut trouver un $\alpha > 0$ tel que $\alpha \leqslant \eta_0$ et

$$\forall s \in [-\alpha, \alpha] \quad \forall t \in [a, b] \quad \left\| \frac{\partial f}{\partial x}(x_0 + s, t) - \frac{\partial f}{\partial x}(x_0, t) \right\| \leqslant \varepsilon$$

On a alors

$$\forall t \in [a, b] \quad \forall s \in [-\alpha, \alpha] \quad \|\varphi_t'(s)\| \leqslant \varepsilon$$

et, par inégalité des accroissements finis

$$\forall t \in [a, b] \quad \forall h \in \mathbb{R} \qquad |h| \leqslant \alpha \Rightarrow \|\varphi_t(h)\| = \|\varphi_t(h) - \varphi_t(0)\| \leqslant \varepsilon |h|$$

Le réel α ne dépend plus de t ! On a alors

$$|h| \leqslant \alpha \Rightarrow \|\Delta(h)\| \leqslant \int_a^b \|\varphi_t(h)\| \, dt \leqslant \varepsilon (b - a) |h|$$

ce qui prouve le résultat, si l'on remarque que le théorème de continuité d'une intégrale dépendant d'un paramètre assure la continuité de $x \mapsto F'(x)$, donc finalement le caractère C^1 de F. ■

REMARQUE 10-2.3 Que penser de l'hypothèse de continuité faite sur f ? A quoi sert-elle ?

EXERCICE 10-2.4 Si, de plus $u : X \to [a, b]$ est de classe C^1, montrer que

$$G(x) = \int_a^{u(x)} f(x, t) \, dt$$

est de classe C^1 sur X, avec

$$G'(x) = u'(x)\ f(x, u(x)) + \int_a^{u(x)} \frac{\partial f}{\partial x}(x,t)\, dt$$

(nous reverrons cet exercice dans le chapitre sur les fonctions de plusieurs variables, mais il peut déjà se traiter "à la main").

EXERCICE 10-2.5 Montrer que

$$F(x) = \int_0^1 \frac{e^{-x^2(1+t^2)}}{(1+t^2)}\, dt$$

définit une fonction de classe C^1 sur $\mathbb{R}$. Si on pose

$$G(x) = \int_0^x e^{-t^2} dt$$

que peut on dire de la fonction $x \mapsto F(x) + G^2(x)$? En déduire la valeur de l'intégrale de Gauss

$$\lim_{x\to+\infty} \int_0^x e^{-t^2} dt$$

EXERCICE 10-2.6 Soit f une fonction C^∞ de $\mathbb{R} \to \mathbb{C}$. On définit $\varphi : \mathbb{R} \to \mathbb{C}$ par

$$\varphi(x) = \frac{f(x) - f(0)}{x}$$

prolongée de manière convenable en 0. Montrer que φ est C^∞ sur $\mathbb{R}$ (mettre φ sous forme intégrale[14]). Généraliser à

$$x \mapsto \frac{f(x) - f(0) - xf'(0) - \cdots - \frac{f^{(n)}(0)}{(n)!}x^n}{x^{n+1}}$$

10-2.4 Théorème de Fubini élémentaire

Le résultat que nous obtenons ici est un cas particulier d'un théorème relatif aux intégrales multiples. Nous le traitons ici dans le cadre des intégrales "simples".

THÉORÈME 10-2.7 Soient $[a,b]$ et $[c,d]$ deux intervalles réels et f une fonction continue $[a,b] \times [c,d] \to E$. On a alors

$$\int_a^b \left(\int_c^d f(x,t)\, dt\right) dx = \int_c^d \left(\int_a^b f(x,t)\, dx\right) dt$$

Démonstration : Remarquons d'abord que les deux membres de l'égalité existent : il résulte en effet du théorème de continuité qui précède que l'application

$$[a,b] \to E \quad x \mapsto \int_c^d f(x,t)\, dt$$

[14]Pour certaines fonctions (comme $x \mapsto \sin x$ par exemple) l'utilisation d'un développement en série entière (voir le chapitre sur les séries de fonctions) simplifie la démonstration.

est continue, et on peut donc parler de son intégrale sur $[a,b]$. Pour prouver l'égalité de ces deux intégrales, nous faisons varier une des bornes d'un des intervalles et considérons les applications

$$H \ : \ [a,b] \to E \quad y \mapsto \int_a^y \left(\int_c^d f(x,t)\, dt \right) dx$$

$$L \ : \ [a,b] \to E \quad y \mapsto \int_c^d \left(\int_a^y f(x,t)\, dx \right) dt$$

Ces deux fonctions vérifient $H(a) = L(a) = 0$. Pour montrer qu'elles sont égales (et $H(b) = L(b)$ est exactement le résultat à démontrer), il suffit de voir que ces deux fonctions sont de classe C^1 et ont la même dérivée.

Pour H, il s'agit de l'intégrale d'une fonction continue, fonction d'une de ses bornes. On a donc bien

$$\forall y \in [a,b] \quad H'(y) = \int_c^d f(y,t)\, dt$$

L se présente comme intégrale à paramètre :

$$L(y) = \int_c^d F(y,t)\, dt \quad \text{avec} \quad F(y,t) = \int_a^y f(x,t)\, dx$$

L'application $F \ : \ [a,b] \times [c,d] \to E$ possède une dérivée partielle par rapport à sa première variable (intégrale d'une fonction continue dépendant d'une de ses bornes, lorsqu'on travaille à t fixé), avec

$$\frac{\partial F}{\partial y}(y,t) = f(y,t)$$

fonction continue sur $[a,b] \times [c,d]$. Pour appliquer le théorème de dérivation vu plus haut, nous vérifions encore que F est continue. Pour (y,t) et $(y_0,t_0) \in [a,b] \times [c,d]$, nous avons

$$\|F(y,t) - F(y_0,t_0)\| \leqslant \left\| \int_a^{y_0} f(x,t)\, dx - \int_a^{y_0} f(x,t_0)\, dx \right\| + \left\| \int_{y_0}^{y} f(x,t)\, dx \right\|$$

et, puisque f est bornée sur le compact $[a,b] \times [c,d]$

$$\|F(y,t) - F(y_0,t_0)\| \leqslant \left\| \int_a^{y_0} f(x,t)\, dx - \int_a^{y_0} f(x,t_0)\, dx \right\| + |y - y_0| \cdot \|f\|_\infty$$

Cette majoration permet de conclure à la continuité de F en (y_0,t_0), en invoquant la continuité de

$$t \mapsto \int_a^{y_0} f(x,t)\, dx$$

On obtient alors

$$L'(y) = \int_c^d \frac{\partial F}{\partial y}(y,t)\,dt = \int_c^d f(y,t)\,dt = H'(y)$$

ce qui démontre le résultat. ■

EXERCICE 10-2.8 Calculer $\int_0^{\pi} \ln(a+\cos t)\,dt$ pour a réel > 1. (Indication : on pourra dériver sous le signe intégrale ou, ce qui revient à la même chose finalement, écrire $\ln(a+\cos t)$ sous forme d'une intégrale).

EXERCICE 10-2.9 Calculer (formellement)

$$\int_0^1 \left(\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\right) dy \quad \text{et} \quad \int_0^1 \left(\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\right) dx$$

Que remarque-t-on ?

10-3 Calculs de primitives

On passe ici en revue différentes techniques de "calcul" de primitives d'une fonction continue $f : I \to \mathbb{C}$ où I est un intervalle de $\mathbb{R}$. La notation

$$F(x) = \int f(x)dx$$

signifie que F est **une** primitive de f sur I, les autres primitives s'en déduisant en ajoutant une constante à F. Il est à noter que, si $x_0 \in I$, la fonction $x \mapsto \int_{x_0}^{x} f(t)dt$ est **la** primitive de f sur I qui s'annule en x_0, mais toute primitive de f n'est pas nécessairement de cette forme puisqu'il n'y a *a priori* aucune raison pour qu'une telle primitive s'annule en au moins un point de I. Penser par exemple à la fonction $x \mapsto \sin x + 2$, primitive de $x \mapsto \cos x$ sur $\mathbb{R}$.

10-3.1 Primitives usuelles

Il s'agit du tableau des dérivées "usuelles", lu à l'envers, avec quelques compléments. Les techniques décrites dans les paragraphes ultérieurs permettent de ré-obtenir certains de ces résultats, par changement de variable notamment.

$f(x)$	$F(x) = \int f(x)dx$	Intervalle de validité
x^α, $\alpha \neq -1$	$\frac{x^{\alpha+1}}{\alpha+1}$	$\mathbb{R}^{*+}$($\mathbb{R}$ si $\alpha \in \mathbb{N}$, $\mathbb{R}^{*-}$aussi si -$\alpha \in \mathbb{N}^*$)
$\frac{1}{x+a}$	$\ln\|x+a\|$	$]-\infty, -a[$ ou $]-a, +\infty[$
$\sin x$	$-\cos x$	$\mathbb{R}$
$\cos x$	$\sin x$	$\mathbb{R}$
$\frac{1}{\cos^2 x}$	$\tan x$	$]-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi[$, avec $k \in \mathbb{Z}$
$\frac{1}{\sin^2 x}$	$-\cot x$	$]k\pi, (k+1)\pi[$ avec $k \in \mathbb{Z}$
$\tan x$	$-\ln\|\cos x\|$	$]-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi[$, avec $k \in \mathbb{Z}$
$\cot x$	$\ln\|\sin x\|$	$]k\pi, (k+1)\pi[$ avec $k \in \mathbb{Z}$
$\frac{1}{\sin x}$	$\ln\left\|\tan\frac{x}{2}\right\|$	$]k\pi, (k+1)\pi[$ avec $k \in \mathbb{Z}$
$\frac{1}{\cos x}$	$\ln\left\|\tan\left(\frac{x}{2}+\frac{\pi}{4}\right)\right\|$	$]-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi[$, avec $k \in \mathbb{Z}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$]-1, 1[$
$\frac{1}{1+x^2}$	$\arctan x$	$\mathbb{R}$

Avec les fonctions exponentielles et hyperboliques :

$f(x)$	$F(x) = \int f(x)dx$	Intervalle de validité
a^x, avec $a > 0$ et $\neq 1$	$\frac{1}{\ln a}a^x$	$\mathbb{R}$
$\operatorname{sh} x$	$\operatorname{ch} x$	$\mathbb{R}$
$\operatorname{ch} x$	$\operatorname{sh} x$	$\mathbb{R}$
$\tanh x$	$\ln \operatorname{ch} x$	$\mathbb{R}$
$\frac{1}{\operatorname{ch}^2 x}$	$\tanh x$	$\mathbb{R}$
$\frac{1}{\operatorname{sh}^2 x}$	$-\coth x$	$]-\infty, 0[$ ou $]0, +\infty[$
$\frac{1}{\operatorname{sh} x}$	$\ln\left\|\tanh\frac{x}{2}\right\|$	$]-\infty, 0[$ ou $]0, +\infty[$
$\frac{1}{\operatorname{ch} x}$	$2\arctan e^x$	$\mathbb{R}$
$\frac{1}{\sqrt{x^2+h}}$, où $h \in \mathbb{R}^*$	$\ln\left\|x+\sqrt{x^2+h}\right\|$	tout intervalle où $x^2+h > 0$

10-3.2 Utilisation de la linéarité

Si f et g sont deux fonctions définies et continues sur I et $\alpha \in \mathbb{C}$ on a évidemment :

$$\int [f(x) + \alpha g(x)]\, dx = \int f(x)dx + \alpha \int g(x)dx$$

(égalité modulo une constante). Pour calculer par exemple

$$F(x) = \int \tan^3(x)dx$$

sur un intervalle de continuité de la fonction $x \mapsto \tan x$, on écrira

$$\int (\tan x(\tan^2 x + 1) - \tan x)dx = \frac{1}{2}\tan^2 x + \ln(|\cos x|)$$

EXERCICE 10-3.1 Trouver ainsi une relation de récurrence permettant de calculer

$$I_n(x) = \int \tan^n x dx, \quad n \in \mathbb{N}$$

On utilise aussi la linéarisation pour calculer les intégrales de la forme

$$I(x) = \int \cos^p x \sin^q x dx$$

où p et q sont deux entiers pairs (dans les autres cas il est préférable d'utiliser un changement de variable). Par exemple pour calculer

$$F(x) = \int (\cos^2 x)(\sin^4 x) dx$$

on écrira

$$\cos^2 x \sin^4 x = \frac{1}{4} (\sin 2x)^2 \sin^2 x = \frac{1}{16}(1 - \cos 4x)(1 - \cos 2x)$$

d'où finalement

$$F(x) = \frac{1}{16} \int \left(1 - \cos 2x - \cos 4x + \frac{1}{2} (\cos 6x + \cos 2x)\right) dx$$
$$= \frac{1}{16} x - \frac{1}{64} \sin 2x - \frac{1}{64} \sin 4x + \frac{1}{192} \sin 6x$$

10-3.3 Intégration par changement de variable

THÉORÈME 10-3.2 Soit $f : I \to \mathbb{C}$ continue et

$$F(x) = \int f(x) dx \quad (1)$$

Si $\varphi : J \to I$ est une application de classe C^1 définie sur un intervalle J à valeurs dans I on a alors

$$H(t) = F(\varphi(t)) = \int f(\varphi(t)) \varphi'(t) dt \quad (2)$$

Démonstration : il s'agit simplement d'appliquer la formule de dérivation d'une composée de fonctions de classe C^1. On remarque que (2) s'obtient "mécaniquement" à partir de (1) en "posant $x = \varphi(t)$", ce qui justifie la notation utilisée pour représenter les primitives. ■

La terminologie "changement de variable" est d'usage, mais s'appliquerait en fait plutôt mieux au corollaire suivant où φ est un difféomorphisme de classe C^1 (donc un véritable changement de variable de classe C^1).

COROLLAIRE 10-3.3 Si $\varphi : J \to I$ est un C^1 difféomorphisme d'intervalles et si $f : I \to \mathbb{C}$ est continue, si

$$H(t) = \int f(\varphi(t)) \varphi'(t) dt \text{ sur } J$$

on a alors

$$\int f(x) dx = H\left(\varphi^{-1}(x)\right) \text{ sur } I$$

EXEMPLE 10-3.4 Pour calculer, sur $]1,+\infty[$, la fonction

$$H(t)=\int\frac{dt}{t\sqrt{t^n-1}}$$

on "posera" $x=\sqrt{t^n-1}$, ce qui revient à considérer l'application C^1 (qui est ici un difféomorphisme)

$$\varphi:]1,+\infty[\to]0,+\infty[\quad t\mapsto\sqrt{t^n-1}$$

On obtient $t^n=x^2+1$, ce qui donne (différentielle logarithmique) $n\frac{dt}{t}=\frac{2xdx}{1+x^2}$ et l'intégrande $\frac{dt}{t\sqrt{t^n-1}}=\frac{dt}{tx}$ se transforme donc en $\frac{2}{n}\frac{dx}{1+x^2}$. Le théorème précédent donne donc

$$H(t)=2\frac{\arctan\sqrt{(t^n-1)}}{n}$$

Remarque : Le changement de variable précédent peut s'obtenir en deux étapes : la "différentielle logarithmique" $\frac{dt}{t}$ est changée en $\frac{1}{n}\frac{du}{u}$ en posant $u=t^n$. Il reste ensuite à calculer $\int\frac{du}{u\sqrt{u-1}}$ qui est du type étudié à la section **10-3.6.3**.

REMARQUE 10-3.5 : *Sur l'homogénéité des formules* : si $(x,\lambda)\mapsto f(x,\lambda)$ est une fonction homogène de degré α -c'est-à-dire que $f(ux,u\lambda)=u^\alpha f(x,\lambda)$- il existe alors une détermination de la primitive

$$F(x,\lambda)=\int f(x,\lambda)\,dx$$

qui est homogène de degré $\alpha+1$. Il suffit en effet, si $G(u)=\int f(u,1)\,du$, de remarquer que, pour $\lambda\neq 0$, on a

$$F(x,\lambda)=\lambda^\alpha\int f\left(\frac{x}{\lambda},1\right)dx=\lambda^{\alpha+1}\int f\left(\frac{x}{\lambda},1\right)d\left(\frac{x}{\lambda}\right)=\lambda^{\alpha+1}G\left(\frac{x}{\lambda}\right)$$

qui est bien homogène de degré $\alpha+1$. Il faut remarquer cependant que l'ajout d'une constante d'intégration (dépendant éventuellement de λ) fait perdre cette homogénéité. Cette remarque peut aider à rectifier certaines erreurs de calcul. Par exemple il n'est pas possible d'avoir

$$\int\frac{dx}{x^2+a^2}=\frac{1}{a^2}\arctan\left(\frac{x}{a}\right),\text{ la bonne valeur étant }\frac{1}{a}\arctan\left(\frac{x}{a}\right)$$

EXERCICE 10-3.6 Calculer $\int\sin^2x\cos^5x\,dx$ (réponse : $\frac{1}{7}\sin^7x-\frac{2}{5}\sin^5x+\frac{1}{3}\sin^3x$).

10-3.4 Intégration par parties

Si f et g sont de classe C^1 sur l'intervalle I, on a

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

Très souvent utilisé pour abaisser le degré d'un polynôme. On peut calculer ainsi les intégrales de la forme

$$\int e^{\lambda x}P(x)dx$$

où P est un polynôme et $\lambda \in \mathbb{C}$. Pour ce type d'intégrale, il est d'ailleurs souvent plus rapide d'utiliser la méthode des coefficients indéterminés, puisqu'un raisonnement élémentaire d'algèbre linéaire assure l'existence d'une primitive de la forme $e^{\lambda x}Q(x)$ où Q est un polynôme de même degré que P.

EXERCICE 10-3.7 Calculer $\int (\cos nx)(x^4+1)\,dx$. La méthode des coefficients indéterminés amène à la recherche d'une primitive de la forme $P(x)\cos nx + Q(x)\sin nx$ où P et Q sont des polynômes de degré au plus 4. Comme la primitive s'annulant en 0 d'une fonction paire est impaire, on cherche la primitive sous la forme

$$(ax^3+bx)\cos nx + (cx^4+dx^2+e)\sin nx$$

Réponse : $(\frac{1}{n}x^4 - \frac{12}{n^3}x^2 + \frac{24}{n^5} + \frac{1}{n})\sin nx + (\frac{4}{n^2}x^3 - \frac{24}{n^4}x)\cos nx$

L'intégration par parties est aussi souvent utilisée pour trouver des relations de récurrence entre primitives. Lorsqu'on cherche à intégrer des fractions rationnelles, on est souvent ramené au calcul de

$$I_n(x) = \int \frac{dx}{(1+x^2)^n}$$

On a alors

$$I_n(x) = \int \frac{1+x^2-x^2}{(1+x^2)^n}dx = I_{n-1}(x) - \int \frac{x}{(1+x^2)^n}xdx$$

Une intégration par parties donne alors

$$\int \frac{x}{(1+x^2)^n}xdx = \frac{1}{2(n-1)}\left[-\frac{x}{(1+x^2)^{n-1}} + \int \frac{dx}{(1+x^2)^{n-1}}\right]$$

ce qui permet d'obtenir une relation de récurrence entre $I_n(x)$ et $I_{n-1}(x)$.

EXERCICE 10-3.8 *Intégrales de Wallis* : trouver une relation de récurrence entre les intégrales

$$J_n(x) = \int \cos^n x\,dx$$

Par un changement de variable, obtenir la relation de récurrence précédemment trouvée entre les intégrales $I_n(x)$.

Enfin l'intégration par parties est un bon moyen de "faire disparaître une fonction transcendante".

EXERCICE 10-3.9 Calculer

$$\int (x^2 + 2x) \arctan x \, dx$$

(réponse $(\frac{1}{3}x^3 + x^2) \arctan x - \frac{1}{6}x^2 - x + \frac{1}{6} \ln\left(1 + x^2\right) + \arctan x$) et

$$\int x \ln(1 + x^2) \, dx$$

(réponse $\frac{1}{2}\left(1 + x^2\right) \ln\left(1 + x^2\right) - \frac{1}{2}x^2$). Sur ce dernier exemple, on voit qu'en posant $g'(x) = x$, les calculs sont plus simples si on prend $g(x) = \frac{1}{2}(x^2 + 1)$.

10-3.5 Intégration des fractions rationnelles

Hormis les cas d'intégration "à vue" (intégrales de la forme $\int \frac{P'(x)}{[P(x)]^n} dx$ notamment), la méthode générale est basée sur une décomposition en éléments simples. On se ramène par division euclidienne à une fraction de degré strictement négatif.

10-3.5.1 Cas d'une décomposition dans $\mathbb{C}(X)$

Si $\alpha \in \mathbb{C}$ est pôle d'ordre p de la fraction rationnelle $R(x)$, la partie polaire de R relative à α est de la forme

$$\sum_{j=1}^{p} \frac{\lambda_j}{(x - \alpha)^j}$$

où les $\lambda_j \in \mathbb{C}$ (avec $\lambda_p \neq 0$). Le coefficient λ_1 est appelé **résidu** de R en α. Toute fraction rationnelle de $\mathbb{C}(X)$ de degré strictement négatif est somme des parties polaires relatives à chacun de ses pôles (ceci est une conséquence du théorème de D'Alembert). On est donc amené à calculer une primitive des éléments simples "de première espèce", soit (à une constante multiplicative près)

$$\int \frac{dx}{(x - \alpha)^j}$$

Si $j \neq 1$ on trouve évidemment

$$\frac{1}{1 - j} \frac{1}{(x - \alpha)^{j-1}}$$

Si $j = 1$ et $\alpha \in \mathbb{R}$, on obtient $\ln |x - \alpha|$, sur un intervalle qui ne contient pas α (un pôle réel de R est évidemment hors du domaine de continuité de R). Lorsque $\alpha = a + ib$, a et b réels avec $b \neq 0$, on obtient, en multipliant haut et bas par le conjugué du dénominateur :

$$\int \frac{dx}{x - (a + ib)} = \int \frac{(x - a)}{(x - a)^2 + b^2} dx + i \int \frac{b}{(x - a)^2 + b^2} dx$$

ce qui donne

$$\int \frac{dx}{x-(a+ib)} = \frac{1}{2}\ln\left((x-a)^2+b^2\right) + i\arctan\left(\frac{x-a}{b}\right)$$

EXERCICE 10-3.10 Montrer qu'une fraction rationnelle R de $\mathbb{C}(X)$ possède une primitive dans $\mathbb{C}(X)$ ssi les résidus relatifs aux différents pôles de R sont tous nuls.

10-3.5.2 Cas d'une décomposition dans $\mathbb{R}(X)$

Si R est à coefficients réels, on peut également travailler avec une décomposition en éléments simples dans $\mathbb{R}(X)$. Si α est un pôle réel de R, l'intégration de la partie polaire relative à α se fait comme précédemment. Si $\alpha \in \mathbb{C} - \mathbb{R}$ est pôle d'ordre m de R, il en est de même de $\overline{\alpha}$, et le trinôme

$$(X-\alpha)(X-\overline{\alpha}) = X^2 + 2pX + q$$

(à discriminant strictement négatif) apparaît avec l'exposant m dans la décomposition en facteurs irréductibles de $\mathbb{R}[X]$ du dénominateur de la fraction R (supposée écrite sous forme irréductible, c'est-à-dire avec un numérateur et un dénominateur premiers entre eux). On obtient alors, dans la décomposition de R, une partie polaire relative à ce trinôme formée d'éléments simples "de deuxième espèce" :

$$\sum_{i=1}^{m} \frac{a_i x + b_i}{(x^2+2px+q)^i}$$

avec $a_i, b_i \in \mathbb{R}$ et $(a_m, b_m) \neq (0,0)$. On a donc à calculer des intégrales de la forme

$$\int \frac{ax+b}{(x^2+2px+q)^n}dx$$

On commence par faire apparaître au numérateur la dérivée du trinôme

$$\int \frac{ax+b}{(x^2+2px+q)^n}dx = \int \frac{a(x+p)}{(x^2+2px+q)^n}dx + \int \frac{b-ap}{(x^2+2px+q)^n}dx$$

La première intégrale se ramène au calcul de $\int \frac{u'(x)}{[u(x)]^n}dx$, qui est élémentaire. Pour calculer la seconde, on met le trinôme sous forme canonique et on obtient

$$\int \frac{dx}{(x^2+2px+q)^n} = \int \frac{dx}{\left[(x+p)^2+r^2\right]^n}$$

avec $r^2 = q - p^2 > 0$. Le changement de variable $x + p = rt$ ramène alors au calcul de

$$I_n(t) = \int \frac{dt}{(1+t^2)^n}$$

dont on a vu le calcul par récurrence à l'aide d'intégrations par parties.

EXERCICE 10-3.11 Calculer

$$\int \frac{x-2}{(x-1)^2(x^2+x+1)^2}dx$$

réponse :

$$\frac{1}{9(x-1)}+\frac{1}{3}\ln(x-1)-\frac{1}{6}\ln\left(x^2+x+1\right)-\frac{5}{9}\sqrt{3}\arctan\frac{\sqrt{3}}{3}(2x+1)-\frac{1}{9}\frac{4x-1}{x^2+x+1}$$

REMARQUE 10-3.12 Dans certains cas, une intégration par parties permet de diminuer le degré du dénominateur avant de se lancer dans une décomposition en éléments simples. Par exemple, pour calculer $\int \frac{dx}{(1-x^3)^2}$ (trois pôles doubles) on écrira

$$\int \frac{dx}{(1-x^3)^2}=\int \frac{1-x^3}{(1-x^3)^2}dx+\int \frac{x^2}{(1-x^3)^2}x\,dx$$

la deuxième intégrale se simplifiant par parties.

EXERCICE 10-3.13 Terminer le calcul précédent. Réponse :

$$-\frac{1}{9(x-1)}-\frac{2}{9}\ln(x-1)+\frac{1}{9}\ln\left(x^2+x+1\right)+\frac{2}{9}\sqrt{3}\arctan\frac{\sqrt{3}}{3}(2x+1)+\frac{1}{9}\frac{x-1}{x^2+x+1}$$

10-3.6 Intégrales se ramenant à des primitives de fractions rationnelles

10-3.6.1 Fonctions rationnelles en sin et cos

Il s'agit de primitives de la forme

$$\int R(\sin x,\cos x)dx$$

où $R(X,Y)=\dfrac{P(X,Y)}{Q(X,Y)}$ est une fraction rationnelle (quotient de deux polynômes) de deux variables. La fonction tan peut intervenir, puisqu'elle est elle-même rationnelle en sin et cos. En particulier les intégrales de la forme

$$\int R(\tan x)dx$$

où $R\in\mathbb{C}(X)$ sont de cette forme.

• Le cas des polynômes $P(\cos x,\sin x)$ a déjà été étudié, le calcul des intégrales $\int \cos^p x\sin^q x\,dx$ distinguant le cas p et q pairs des autres cas.

• Sur un intervalle de continuité de la fonction $F(x)=R(\sin x,\cos x)$ inclus dans un intervalle de la forme $I_k=](2k-1)\pi,(2k+1)\pi[$ avec $k\in\mathbb{Z}$, on peut faire le changement de variable $t=\tan\frac{x}{2}$, ce qui correspond au C^∞-difféomorphime

$$I_k\to\mathbb{R}\quad x\mapsto\varphi(x)=\tan\frac{x}{2}=t,\text{ avec }\varphi^{-1}(t)=2(\arctan t+k\pi)$$

d'où $dx = \dfrac{2dt}{1+t^2}$ et donc, sur un intervalle convenable **inclus dans** I_k

$$F(x) = \int R(\sin x, \cos x)dx = H\left(\tan\frac{x}{2}\right)$$

$$\text{avec } H(t) = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}$$

EXEMPLE 10-3.14 La fonction $f : x \mapsto \dfrac{1}{4\sin x - 5\cos x + 7}$ est 2π-périodique et continue sur $\mathbb{R}$ (puisque la fonction $x \mapsto 4\sin x - 5\cos x$ peut s'écrire sous la forme $\sqrt{4^2+5^2}\sin(x-\alpha)$ et est donc inférieure à 7 en valeur absolue). Elle possède donc des primitives sur $\mathbb{R}$. Sur tout intervalle de la forme I_k, la démarche précédente donne

$$\int \frac{dx}{4\sin x - 5\cos x + 7} = \int \frac{dt}{6t^2+4t+1} = \frac{1}{\sqrt{2}}\arctan\sqrt{2}\left(3\tan\left(\frac{x}{2}\right)+1\right)$$

Attention ! Si on note $G(x)$ l'expression précédente, on a $G(2\pi)-G(0) = 0$ et ce serait une lourde faute d'en déduire que $\displaystyle\int_0^{2\pi}\frac{dx}{4\sin x - 5\cos x + 7} = 0$, puisque la fonction f est continue et strictement positive. G n'est pas l'expression d'une primitive de f sur $[0, 2\pi]$. L'intégrale de f sur une période est $G(\pi^-)-G(-\pi^+) = \dfrac{\pi}{\sqrt{2}}$ et la valeur moyenne de f est donc $\dfrac{1}{2\sqrt{2}}$. On peut d'ailleurs en déduire l'expression d'une primitive de f qui soit valable sur $\mathbb{R}$. La fonction $f - \dfrac{1}{2\sqrt{2}}$ est 2π-périodique, de valeur moyenne nulle et possède donc des primitives périodiques. Si F est une primitive de f sur $\mathbb{R}$, la fonction $x \mapsto F(x) - \dfrac{x}{2\sqrt{2}}$ est donc 2π-périodique. Sur $]-\pi,\pi[$, on peut prendre l'expression

$$\frac{1}{\sqrt{2}}\arctan\sqrt{2}\left(3\tan\left(\frac{x}{2}\right)+1\right) - \frac{x}{2\sqrt{2}}$$
$$= \frac{1}{\sqrt{2}}\left[\arctan\sqrt{2}\left(3\tan\left(\frac{x}{2}\right)+1\right) - \arctan\left(\tan\left(\frac{x}{2}\right)\right)\right]$$

Si $ab \neq -1$, il est facile de voir que $\arctan a - \arctan b$ et $\arctan\dfrac{a-b}{1+ab}$ diffèrent d'un multiple de π. Comme une fonction continue sur un intervalle à valeurs dans $\pi\mathbb{Z}$ est forcément constante (théorème des valeurs intermédiaires) , on peut prendre comme expression sur $]-\pi,\pi[$

$$F(x) - \frac{x}{2\sqrt{2}} = \frac{1}{\sqrt{2}}\arctan\left(\frac{(3\sqrt{2}-1)\tan\frac{x}{2}+\sqrt{2}}{1+\sqrt{2}\left(3\tan\frac{x}{2}+1\right)\tan\frac{x}{2}}\right)$$
$$= \frac{1}{2}\sqrt{2}\arctan\left(\frac{\sqrt{2}+\sqrt{2}\cos x + (3\sqrt{2}-1)\sin x}{1+3\sqrt{2}+(1-3\sqrt{2})\cos x + \sqrt{2}\sin x}\right)$$

Cette fonction est continue sur $\mathbb{R}$ (vérification facile) et est clairement 2π-périodique. On peut donc prendre

$$\forall\, x \in \mathbb{R} \quad F(x) = \frac{x}{2\sqrt{2}} + \frac{1}{\sqrt{2}} \arctan\left(\frac{\sqrt{2} + \sqrt{2}\cos x + (3\sqrt{2} - 1)\sin x}{1 + 3\sqrt{2} + (1 - 3\sqrt{2})\cos x + \sqrt{2}\sin x}\right)$$

Il est parfois plus simple d'utiliser d'autres changements de variables :

- Lorsque "l'intégrande" $R(\sin x, \cos x)\,dx$ peut s'écrire (après manipulations algébriques) $R_1(\cos x)\sin x\,dx$, où R_1 est une fraction rationnelle à une indéterminée, on voit que le changement de variable $u = \cos x$ amène au calcul de

  $$-\int R_1(u)\,du$$

  Une condition nécessaire pour qu'il en soit ainsi est que l'intégrande soit invariant par changement de x en $-x$, puisque

  $$R_1(\cos(-x))\sin(-x)\,d(-x) = R_1(\cos x)\sin x\,dx$$

  On peut montrer (voir plus loin) que cette condition est aussi suffisante.

- Lorsque l'intégrande peut se mettre sous la forme $R_1(\sin x)\cos x\,dx$ (une condition nécessaire et suffisante pour qu'il en soit ainsi est que l'intégrande soit invariant en changeant x en $\pi - x$), on utilisera de même le changement de variable $u = \sin x$.

- Enfin lorsque l'intégrande est invariant par changement de x en $\pi + x$, on montre qu'il peut s'écrire $R_1(\tan x)\,dx$ ou aussi $\dfrac{R_1(\tan x)}{1 + \tan^2 x}(1 + \tan^2 x)\,dx$ et le changement de variable $u = \tan x$ (valable dans tout intervalle de la forme $]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$ avec $k \in \mathbb{Z}$) peut être utilisé.

EXEMPLE 10-3.15 Calculer $\displaystyle\int \frac{\cos^2 x}{\sin 2x(1 + \tan x)}dx$. On vérifie ici que l'intégrande est invariant en changeant x en $\pi + x$. Le changement de variable $u = \tan x$ est donc possible :

$$\frac{\cos^2 x}{\sin 2x(1 + \tan x)} = \frac{\cos^2 x}{2\sin x\cos x(1 + \tan x)(1 + \tan^2 x)}(1 + \tan^2 x)$$

d'où

$$\int \frac{\cos^2 x}{(\sin 2x)(1 + \tan x)}dx = \int \frac{du}{2u(1 + u)(1 + u^2)}$$
$$= -\frac{1}{8}\ln(u^2 + 1) - \frac{1}{4}\arctan u + \frac{1}{2}\ln u - \frac{1}{4}\ln(1 + u)$$

soit enfin

$$\int \frac{\cos^2 x}{(\sin 2x)(1 + \tan x)}dx = -\frac{1}{4}x + \frac{1}{2}\ln|\sin x| - \frac{1}{4}\ln|\cos x + \sin x|$$

Justifions par exemple le changement de variable $u = \cos x$ lorsque l'intégrande $R(\cos x, \sin x)\, dx$ est formellement invariante en changeant x en $-x$. Ceci signifie que

$$R(X, -Y) = -R(X, Y)$$

La fraction rationnelle $R(X, Y)$ peut s'écrire, en séparant au numérateur et au dénominateur les termes pairs et impairs en Y

$$R(X,Y) = \frac{A(X,Y^2) + YB(X,Y^2)}{A_1(X,Y^2) + YB_1(X,Y^2)} = \frac{P_1(X,Y^2) + YP_2(X,Y^2)}{Q(X,Y^2)}$$

en multipliant haut et bas par $A_1(X,Y^2) - YB_1(X,Y^2)$. La condition d'imparité $R(X,-Y) = -R(X,Y)$ donne alors $P_1 = 0$ et

$$R(\cos x, \sin x)\, dx = \frac{P_2(\cos x, 1 - \cos^2 x)}{Q(\cos x, 1 - \cos^2 x)} \sin x\, dx = R_1(\cos x) \sin x\, dx$$

où R_1 est une fraction rationnelle, écriture qui permet le changement de variable $u = \cos x$. On justifierait de manière analogue (exercice) les changements de variables correspondant aux autres propriétés d'invariance de l'intégrande.

Remarques :

- Lorsque l'intégrande est invariant par les trois changements vus précédemment, il peut parfois se mettre sous la forme $R_1(\cos 2x) \sin 2x\, dx$, d'où le changement de variable $u = \cos 2x$ (ou de manière équivalente $\cos^2 x$). Calculer par exemple

$$\int \frac{\tan x}{2\tan^2 x + 4\cos^2 x + 6} dx$$

On fait apparaître l'expression $\sin x \cos x dx$ en multipliant haut et bas par $\cos^2 x$. On obtient

$$\int \frac{\sin x \cos x}{4\cos^4 x + 4\cos^2 x + 2} dx$$

et finalement

$$\int \frac{\tan x}{2\tan^2 x + 4\cos^2 x + 6} dx = -\frac{1}{4} \arctan\left(2\cos^2 x + 1\right)$$

- Ne pas oublier que $\cos nx$ et $\sin nx$ ($n \in \mathbb{N}^*$) sont des polynômes en sin et cos. Calculer par exemple

$$\int \frac{\sin\left(\frac{x}{2}\right)}{\cos\left(\frac{x}{3}\right)} dx$$

(On utilise la variable $u = \dfrac{x}{6}$ puis -propriété d'invariance de l'intégrande- $v = \cos u$, et donc faire apparaître $\sin u\, du$ en écrivant

$$\sin 3u = \sin u(3 - 4\sin^2 u) = \sin u(4\cos^2 u - 1)$$

Réponse :

$$-12\cos\frac{x}{6}+\frac{3\sqrt{2}}{2}\ln\left|\frac{\sqrt{2}\cos\frac{x}{6}+1}{\sqrt{2}\cos\frac{x}{6}-1}\right|$$

- Avant de faire les changements de variable, il faut commencer par simplifier les fractions, notamment en faisant apparaître les parties entières. Par exemple, pour calculer

$$\int\frac{\sin 3x}{2\sin x+1}dx$$

il n'y a pas de propriété d'invariance particulière de l'intégrande. On utilisera donc le changement de variable $u=\tan\frac{x}{2}$ mais on écrira d'abord

$$\frac{\sin 3x}{2\sin x+1}=\frac{3\sin x-4\sin^3 x}{2\sin x+1}=-2\sin^2 x+\sin x+1-\frac{1}{2\sin x+1}$$

(division euclidienne) et on en déduit

$$\int\frac{\sin 3x}{2\sin x+1}dx=\sin x\cos x-\cos x+\frac{\sqrt{3}}{3}\ln\left|\frac{\frac{\sqrt{3}}{3}\left(\tan\frac{x}{2}+2\right)+1}{\frac{\sqrt{3}}{3}\left(\tan\frac{x}{2}+2\right)-1}\right|$$

- Noter enfin que les changements de variable sensés simplifier les calculs peuvent parfois les compliquer. Pour calculer

$$\int\frac{dx}{\sin^3 x}$$

$u=\cos x$ est un changement de variable plus compliqué que $\tan\frac{x}{2}$ (pourquoi ?). On obtient

$$\int\frac{dx}{\sin^3 x}=\frac{1}{8}\tan^2\frac{x}{2}+\frac{1}{2}\ln\left(\tan\frac{x}{2}\right)-\frac{1}{8\tan^2\frac{x}{2}}$$

10-3.6.2 Fractions rationnelles en e^x, sh$x$ et ch$x$

- Pour calculer les primitives de la forme $\int R(\operatorname{ch}x,\operatorname{sh}x)\,dx$, on utilise en général le changement de variable hyperbolique correspondant au changement de variable trigonométrique qu'on utiliserait pour calculer $\int R(\cos x,\sin x)\,dx$. Ceci est dû au fait que les propriétés "algébriques" des fonctions hyperboliques sont pratiquement les mêmes que celles des fonctions circulaires. Rappelons à ce propos que les formules de trigonométrie hyperboliques peuvent se déduire de celles de

trigonométrie circulaire (ne faisant pas intervenir la 2π-périodicité évidemment !) en remplaçant cos par ch, sin par i sh (et tan par i tanh). Par exemple

$$\int \frac{dx}{\mathrm{ch}^4 x}$$

se calculera à l'aide du changement de variable $u = \tanh x$ puisque $\dfrac{dx}{\cos^4 x}$ est invariant en changeant x en $\pi + x$. On obtient :

$$\int \frac{dx}{\mathrm{ch}^4 x} = \tanh x - \frac{1}{3}\tanh^3 x$$

• Comme $\mathrm{sh}\, x$ et $\mathrm{ch}\, x$ sont des fractions rationnelles en e^x, les primitives précédentes sont toutes de la forme $\int R_1(e^x)\, dx$, où R_1 est une fraction rationnelle. On dispose aussi pour traiter ce genre de primitive du changement de variable $u = e^x$, qui ramène à la primitive $\displaystyle\int \frac{R_1(u)}{u} du$. Calculer par exemple

$$\int \frac{\mathrm{sh}\, x}{e^{2x} + e^x + 1} dx$$

Réponse : $\dfrac{e^{-x}}{2} + \dfrac{x}{2} - \dfrac{1}{4}\ln\left(e^{2x} + e^x + 1\right) + \dfrac{\sqrt{3}}{2}\arctan\dfrac{\sqrt{3}}{3}(2e^x + 1)$

10-3.6.3 Intégrales abéliennes

Il s'agit ici pour nous d'intégrales de la forme

$$\int R(x, \varphi(x))\, dx$$

où R est une fraction rationnelle et où φ est de la forme

$$\varphi(x) = \sqrt[n]{\left(\frac{ax+b}{cx+d}\right)} \quad \text{ou} \quad \varphi(x) = \sqrt{ax^2 + bx + c}$$

remarquons que dans le premier cas, la fonction homographique peut être une fonction affine ($c = 0$ et $d = 1$).

Intégrales $\displaystyle\int R\left(x, \sqrt[n]{\left(\frac{ax+b}{cx+d}\right)}\right) dx$

On utilise le changement de variable $u = \sqrt[n]{\left(\dfrac{ax+b}{cx+d}\right)}$, x est donc fonction rationnelle de u et on arrive à une primitive d'une fonction rationnelle en u. Calculer par exemple

$$\int \sqrt[3]{\frac{x-1}{x+1}} dx$$

Le changement de variable $u = \sqrt[3]{\dfrac{x-1}{x+1}}$, $x = \dfrac{1+u^3}{1-u^3}$ amène à

$$\int 6\frac{u^3}{(u^3-1)^2}\,du = -\frac{2u}{u^3-1} + \int \frac{2}{u^3-1}\,du$$

par parties soit enfin

$$-2\frac{u}{u^3-1} + \frac{2}{3}\ln(u-1) - \frac{1}{3}\ln\left(u^2+u+1\right) - \frac{2\sqrt{3}}{3}\arctan\frac{\sqrt{3}}{3}(2u+1)$$

et donc

$$\int \sqrt[3]{\frac{x-1}{x+1}}dx = (x+1)\sqrt[3]{\frac{x-1}{x+1}} + \ln\left|\sqrt[3]{\left(\frac{x-1}{x+1}\right)} - 1\right|$$
$$+\frac{1}{3}\ln|x+1| - \frac{2\sqrt{3}}{3}\arctan\frac{1}{3}\left(2\sqrt[3]{\left(\frac{x-1}{x+1}\right)} + 1\right)$$

Intégrales de la forme $\int R\left(x, \sqrt{ax^2+bx+c}\right)\,dx$

• Lorsque les coefficients a, b, c sont fixés, on met en général le trinôme du second degré $ax^2 + bx + c$ sous forme canonique et un changement de variable affine ramène à une des trois formes

$$\int R_1(u, \sqrt{1+u^2})du \text{ ou } \int R_1(u, \sqrt{1-u^2})du \text{ ou } \int R_1(u, \sqrt{u^2-1})du$$

avec R_1 fraction rationnelle. Dans le premier cas un changement de variable $u = \operatorname{sh} t$ (parfois $u = \tan t$) ramène à un type étudié plus haut. Dans le second cas $u = \sin t$, soit $t = \arcsin u$, et dans le troisième $u = \pm\operatorname{ch} t$ (parfois $u = \dfrac{1}{\cos t}$) donnent des résultats analogues.

EXEMPLE 10-3.16 Calculer $\int x\sqrt{x^2+x+1}dx$. On écrit

$$x^2+x+1 = \left(x+\frac{1}{2}\right)^2 + \frac{3}{4}$$

et on fait donc le changement de variable $x + \dfrac{1}{2} = \dfrac{\sqrt{3}}{2}\operatorname{sh} t$. On obtient

$$\frac{3}{4}\int\left(\frac{\sqrt{3}}{2}\operatorname{sh} t - \frac{1}{2}\right)\operatorname{ch}^2 t\,dt$$

soit finalement

$$\frac{3}{8}\sqrt{3}\left(\frac{\operatorname{ch}^3 t}{3}\right) - \frac{3}{16}\operatorname{ch} t\operatorname{sh} t - \frac{3}{16}t$$

On a donc

$$\int x\sqrt{x^2+x+1}dx = \frac{1}{3}\left(\sqrt{(x^2+x+1)}\right)^3$$
$$-\frac{1}{8}(2x+1)\sqrt{(x^2+x+1)} - \frac{3}{16}\text{argsh}\frac{2\sqrt{3}}{3}\left(x+\frac{1}{2}\right)$$

• On peut parfois aussi utiliser un paramétrage rationnel de la demi-conique Γ d'équation

$$y = \sqrt{ax^2+bx+c}$$

Si on obtient un tel paramétrage

$$\begin{cases} x &= r_1(t) \\ y &= r_2(t) \end{cases}$$

avec r_1 et r_2 fonctions rationnelles, on tombera sur

$$\int R\left(x, \sqrt{ax^2+bx+c}\right)dx = \int R(r_1(t), r_2(t))\, r_1'(t)\, dt$$

soit un calcul de primitive de fraction rationnelle. On remarquera que cette méthode permet de ramener à un calcul de primitive de fonction rationnelle toute intégrale de la forme

$$\int R(x, f(x))dx$$

où R est rationnelle et où la courbe d'équation $y = f(x)$ admet un paramétrage rationnel. C'est d'ailleurs cette démarche qui a été suivie dans le paragraphe précédent.

Pour obtenir un paramétrage rationnel d'une conique, on remarque qu'une droite du plan coupe "en général" la conique en deux points. Si on astreint une droite variable $\mathcal{D}_t$ à passer par un point fixe de la conique, il y a un autre point d'intersection dont les coordonnées sont fonctions rationnelles de la pente t de $\mathcal{D}_t$.

- Si le trinôme ax^2+bx+c possède des racines (la conique Γ d'équation $y^2 = ax^2+bx+c$, est du genre hyperbole d'axe focal Ox si $a > 0$, du genre ellipse si $a < 0$), il se factorise en

  $$ax^2+bx+c = a(x-\alpha)(x-\beta)$$

  On fait en général passer la droite $\mathcal{D}_t$ par un des sommets soit $A(\alpha, 0)$. Elle a alors pour équation

  $$(\mathcal{D}_t) \quad y = t(x-\alpha)$$

  L'abscisse de l'autre point commun à Γ et $\mathcal{D}_t$ vérifie donc

  $$[t(x-\alpha)]^2 = a(x-\alpha)(x-\beta)$$

  soit $x(t) = \dfrac{(t^2\alpha - a\beta)}{(t^2-a)}$ et $y(t) = t(x(t)-\alpha) = \dfrac{(\alpha-\beta)ta}{t^2-a}$, ce qui donne bien un paramétrage rationnel de Γ (figure **10.1**).

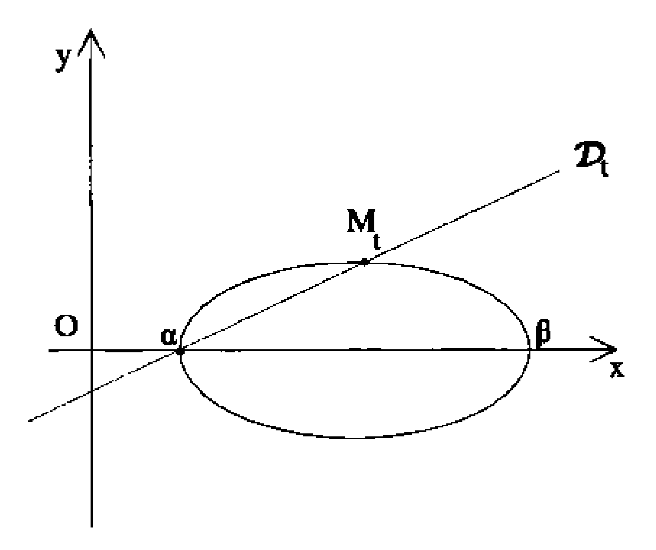


Figure **10.1** – *Paramétrage rationnel de l'ellipse*

- Si le trinôme n'a pas de racine, il est du signe de a et étant sous le radical, on doit avoir $a > 0$. La conique Γ est alors une hyperbole d'axe focal parallèle à Oy. On peut alors (et la méthode est valable aussi lorsque Ox est axe focal) couper Γ par une droite $\mathcal{D}'_t$ parallèle à une asymptote, donc d'équation

$$(\mathcal{D}'_t) \quad y = \sqrt{a}x + t$$

$\mathcal{D}'_t$ et Γ ont un point d'intersection rejeté à l'infini, l'abscisse de l'autre point d'intersection vérifiant $(\sqrt{a}x + t)^2 = ax^2 + bx + c$, soit

$$x(t) = \frac{c - t^2}{2\sqrt{a}t - b} \quad \text{et} \quad y(t) = \sqrt{a}x(t) + t = \frac{\sqrt{a}t^2 - tb + \sqrt{a}c}{2\sqrt{a}t - b}$$

EXEMPLE 10-3.17 Calculer

$$\int \frac{dx}{x\sqrt{(x-1)(4x-2)}}$$

On coupe la conique par la droite d'équation $y = t(x-1)$. On obtient $x = \dfrac{t^2-2}{t^2-4}$ et l'intégrale $-\displaystyle\int \frac{2dt}{t^2-2} = -\frac{\sqrt{2}}{2}\ln\left|t-\sqrt{2}\right| + \frac{\sqrt{2}}{2}\ln\left|t+\sqrt{2}\right|$. Il reste ensuite à remplacer t par son expression $t = \dfrac{y}{x-1} = \dfrac{\sqrt{(x-1)(4x-2)}}{x-1}$.

EXERCICE 10-3.18 Calculer $\displaystyle\int \sqrt[3]{x^3 - x^2 - x + 1}dx$

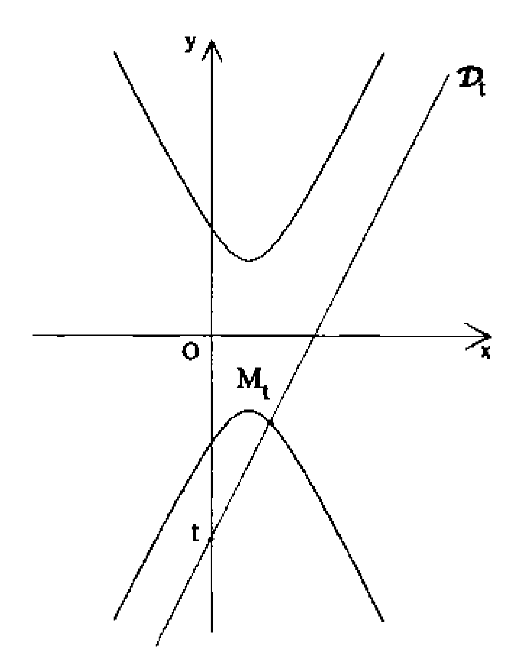


Figure **10.2** – *Paramétrage rationnel de l'hyperbole*

10-4 Calculs approchés d'intégrales

10-4.1 Méthode des trapèzes

On utilise la formule d'intégration approchée

$$\int_\alpha^\beta f(t)dt \simeq (\beta-\alpha)\frac{f(\alpha)+f(\beta)}{2}$$

formule exacte si f est polynomiale de degré inférieur ou égal à 1. En décomposant ensuite le segment $[a,b]$ en n segments de longueur $\frac{b-a}{n}$, et en appliquant sur chaque sous-segment l'égalité approchée précédente on obtient

$$\int_a^b f(t)dt \simeq T_n(f,a,b) = \frac{b-a}{n}\left[\frac{1}{2}f(a)+\sum_{k=1}^{n-1} f\left(a+k\frac{b-a}{n}\right)+\frac{1}{2}f(b)\right]$$

Majoration de l'erreur : On suppose ici la fonction de classe C^2 sur le segment $[a,b]$, (l'existence d'une dérivée seconde non nulle mesurant d'une certaine façon ce qui différencie f d'une fonction polynomiale de degré 1 et donc ce qui empêche la formule d'intégration précédente d'être exacte) et on majore d'abord chacune des erreurs

$$\Delta(\alpha,\beta) = \left|\int_\alpha^\beta f(t)dt - (\beta-\alpha)\frac{f(\alpha)+f(\beta)}{2}\right|$$

$[\alpha,\beta]$ représentant un des intervalles de la subdivision de $[a,b]$ considérée précédemment. On note P le polynôme de degré inférieur ou égal à 1 qui vérifie

$$P(\alpha)=f(\alpha) \text{ et } P(\beta)=f(\beta)$$

Comme la formule d'intégration est exacte pour le polynôme P, on a

$$\Delta(\alpha,\beta) = \left|\int_\alpha^\beta f(t)dt - \int_\alpha^\beta P(t)dt\right| \leqslant \int_\alpha^\beta |f(t) - P(t)|\, dt$$

Si f est polynomiale de degré 2 il en est de même de $f - P$, fonction qui s'annule en α et β et est donc de la forme $f(t) - P(t) = (t-\alpha)(t-\beta)\frac{A}{2}$ où A est la valeur (constante) de la dérivée f''. Dans le cas général, si $x \in]\alpha,\beta[$, on choisit le réel λ pour avoir $f(x) - P(x) = (x-\alpha)(x-\beta)\frac{\lambda}{2}$. La fonction de classe C^2 $t \mapsto f(t) - P(t) - (t-\alpha)(t-\beta)\frac{\lambda}{2}$ s'annule en α, β et x. En appliquant 2 fois le théorème de Rolle (on suppose que f est à valeurs réelles) on obtient l'existence d'un réel $\gamma \in [\alpha,\beta]$ tel que $\lambda = f''(\gamma)$. On en déduit la majoration

$$\forall\, x \in [\alpha,\beta] \quad |f(x) - P(x)| \leqslant \frac{(x-\alpha)(\beta-x)}{2}\, \|f''\|_\infty$$

faisant intervenir la borne supérieure de $|f''|$ sur $[\alpha,\beta]$, qu'on peut encore majorer par la borne supérieure sur $[a,b]$. On obtient finalement

$$\Delta(\alpha,\beta) \leqslant \int_\alpha^\beta |f(t) - P(t)|\, dt \leqslant \int_\alpha^\beta \frac{(x-\alpha)(\beta-x)}{2}\, \|f''\|_\infty\, dx = \frac{1}{12}(\beta-\alpha)^3\, \|f''\|_\infty$$

et donc dans le cas de l'utilisation de n points de subdivision, après majoration des n erreurs élémentaires on a

$$\left|\int_a^b f(t)dt - T_n(f,a,b)\right| \leqslant \frac{(b-a)^3}{12n^2}\, \|f''\|_\infty$$

10-4.2 Méthode de Simpson

On utilise cette fois la formule d'intégration approchée

$$\int_\alpha^\beta f(t)dt \simeq \frac{(\beta-\alpha)}{6}\left[f(\alpha) + f(\beta) + 4f\left(\frac{\alpha+\beta}{2}\right)\right]$$

formule dont on vérifie aisément qu'elle est exacte si f est polynomiale de degré inférieur ou égal à 3. En décomposant ensuite le segment $[a,b]$ en n segments de longueur $\frac{b-a}{n}$, et en appliquant sur chaque sous-segment l'égalité approchée précédente on obtient

$$\int_a^b f(t)dt \simeq S_n(f,a,b) = \frac{b-a}{6n}\left[f(a) + 2\sum_{k=1}^{n-1} f(a_k) + 4\sum_{k=0}^{n-1} f(b_k) + f(b)\right]$$

avec $a_k = a + k\frac{b-a}{n}$ et $b_k = \frac{a_k + a_{k+1}}{2}$.

Majoration de l'erreur : On suppose ici la fonction de classe C^4 sur le segment $[a,b]$ et on majore d'abord chacune des erreurs "élémentaire"

$$\Delta(\alpha,\beta) = \left|\int_\alpha^\beta f(t)dt - \frac{(\beta-\alpha)}{6}\left[f(\alpha) + f(\beta) + 4f\left(\frac{\alpha+\beta}{2}\right)\right]\right|$$

$[\alpha, \beta]$ représentant un des intervalles de la subdivision de $[a, b]$ considérée précédemment. Montrer qu'il existe un unique polynôme P de degré inférieur ou égal à 3 qui vérifie

$$P(\alpha) = f(\alpha),\ P(\beta) = f(\beta),$$
$$P\left(\frac{\alpha+\beta}{2}\right) = f\left(\frac{\alpha+\beta}{2}\right) \text{ et}$$
$$P'\left(\frac{\alpha+\beta}{2}\right) = f'\left(\frac{\alpha+\beta}{2}\right)$$

Comme la formule d'intégration est exacte pour le polynôme P, on a

$$\Delta(\alpha, \beta) = \left|\int_\alpha^\beta f(t)dt - \int_\alpha^\beta P(t)dt\right| \leqslant \int_\alpha^\beta |f(t) - P(t)|\, dt$$

Montrer la majoration

$$\forall\, x \in [\alpha, \beta] \quad |f(x) - P(x)| \leqslant \frac{(x-\alpha)(\beta - x)\left(x - \dfrac{\alpha+\beta}{2}\right)^2}{24} \left\|f^{(4)}\right\|_\infty$$

et en déduire

$$\Delta(\alpha, \beta) \leqslant \frac{1}{2880} (\beta - \alpha)^5 \left\|f^{(4)}\right\|_\infty$$

Dans le cas de l'utilisation de n points de subdivision, après majoration des n erreurs élémentaires on a

$$\left|\int_a^b f(t)dt - S_n(f, a, b)\right| \leqslant \frac{(b-a)^5}{2880 n^4} \left\|f^{(4)}\right\|_\infty$$

10-4.3 Accélération de convergence : méthode de Romberg

10-4.3.1 Procédé d'extrapolation de Richardson

On suppose que l'on étudie une fonction A définie au voisinage de 0 et admettant un développement limité à tout ordre

$$A(t) = a_0 + a_1 t + \cdots + a_p t^p + O(t^{p+1})$$

La valeur a_0 de A en zéro n'est pas connue, mais l'on dispose d'un algorithme permettant de calculer des valeurs

$$\alpha_m = A(t_m) \text{ avec } t_m \to 0^+$$

formant une suite qui converge vers a_0. On conçoit que, plus l'ordre de l'infiniment petit (en 0) $\delta(t) = A(t) - a_0$ est élevé, plus la suite (α_m) convergera rapidement vers a_0. Le principe de la méthode de Richardson est le suivant :

On choisit un réel fixe $r > 1$. On pose $A_0(t) = A(t)$ et on a alors

$$A_0(rt) = a_0 + a_1 rt + \cdots + a_p r^p t^p + O(t^{p+1})$$

Pour faire disparaître le terme en t sans changer la valeur a_0, on voit qu'il suffit de poser

$$A_1(t) = \frac{rA_0(t) - A_0(rt)}{r-1}$$

et on a alors

$$A_1(t) = a_0 + O(t^2)$$

En définissant ensuite, de proche en proche

$$A_2(t) = \frac{r^2 A_1(t) - A_1(rt)}{r^2 - 1} \quad \cdots \quad A_k(t) = \frac{r^k A_{k-1}(t) - A_{k-1}(rt)}{r^k - 1}$$

on aura

$$A_k(t) = a_0 + O(t^{k+1})$$

Pour "t petit", $A_k(t)$ est donc meilleure approximation de a_0 que $A_1(t)$.

On obtient un procédé d'accélération de convergence lorsqu'on part de la suite

$$\alpha_m = A_{m,0} = A(\frac{t_0}{r^m})$$

qui converge vers a_0 puisqu'on a choisi $r > 1$ (t_0 est fixé) et vérifie

$$A_{m,0} \underset{m\to\infty}{=} a_0 + O(r^{-m})$$

La suite

$$A_{m,1} = A_1(\frac{t_0}{r^m})$$

vérifie $A_{m,1} \underset{m\to\infty}{=} a_0 + O(r^{-2m})$ (et converge donc grosso modo deux fois plus vite) et de même

$$A_{m,k} = A_k(\frac{t_0}{r^m})$$

vérifie $A_{m,k} \underset{m\to\infty}{=} a_0 + O(r^{-(k+1)m})$ et converge $k+1$ fois plus vite que la suite de départ. Les définitions précédentes montrent que les termes de ces différentes suites se calculent de proche en proche

$$A_{m,k} = \frac{r^k A_{m,k-1} - A_{m-1,k-1}}{r^k - 1}$$

Pratiquement, on présente les calculs sous forme d'un tableau

$$\begin{array}{ccccccc}
A_{0,0} & \to & A_{1,1} & \to & A_{2,2} & \to & A_{3,3} \\
 & \nearrow & & \nearrow & & \nearrow & \\
A_{1,0} & \to & A_{2,1} & \to & A_{3,2} & \to & A_{3,4} \\
 & \nearrow & & \nearrow & & \nearrow & \\
A_{2,0} & \to & A_{3,1} & \to & A_{4,2} & & \\
 & \nearrow & & \nearrow & & & \\
A_{3,0} & \to & A_{4,1} & \to & & & \\
 & \nearrow & & & & &
\end{array}$$

chaque colonne du tableau formant une suite convergeant vers a_0 avec gain de rapidité lorsqu'on se déplace vers la droite.

10-4.3.2 Méthode de Romberg

Si f est une fonction continue sur $[a,b]$, et si

$$T_n(f,a,b) = \frac{b-a}{n}\left[\frac{1}{2}f(a) + \sum_{k=1}^{n-1} f\left(a + k\frac{b-a}{n}\right) + \frac{1}{2}f(b)\right]$$

est la valeur approchée de son intégrale sur $[a,b]$ par la méthode des trapèzes, un argument de continuité uniforme montre que

$$\lim_{n\to+\infty} T_n(f,a,b) = \int_a^b f(t)dt$$

La méthode de Romberg correspond à une accélération de cette convergence basée sur le procédé de Richardson, en ajoutant quelques hypothèses de régularité pour la fonction.

Si f est de classe C^{2k+2} sur $[a,b]$, on démontre que $T_n(f,a,b)$ possède un développement limité pour $n \to +\infty$ d'ordre $2k+1$ de la forme

$$T_n(f,a,b) = \int_a^b f(t)dt + \frac{c_1}{n^2} + \frac{c_2}{n^4} + \cdots + \frac{c_k}{n^{2k}} + O(\frac{1}{n^{2k+2}})$$

On note h la variable $\dfrac{b-a}{n}$ et on a alors

$$T_n(f,a,b) = T(h) = \int_a^b f(t)dt + a_1h^2 + \cdots + a_kh^{2k} + O(h^{2k+2})$$

En posant $T(h) = A(h^2)$, la fonction A possède au voisinage de zéro à droite un développement limité d'ordre k

$$A(h) = \int_a^b f(t)dt + a_1h + \cdots + a_kh^k + O(h^{k+1})$$

On applique alors le procédé d'extrapolation de Richardson avec

$$A_{m,0} = A\left(\frac{(b-a)^2}{4^m}\right)$$

et donc $r = 4$, ce qui correspond à une formule des trapèzes avec 2^m points de subdivision (l'intérêt de multiplier par 2 le nombre de points de subdivision à chaque étape est que le calcul de $A_{m-1,0}$ est utilisable pour obtenir $A_{m,0}$, il ne faut évaluer la fonction f qu'aux nouveaux points intermédiaires de subdivision). Les suites $(A_{m,p})_{m\in\mathbb{N}-\{0,\dots,p-1\}}$ déterminées par

$$A_{m,p} = \frac{4^pA_{m,p-1} - A_{m-1,p-1}}{4^p - 1}$$

convergent d'autant plus rapidement que p est grand.

EXEMPLE 10-4.1 Testons la méthode sur un exemple simple. La formule des trapèzes pour obtenir une valeur approchée de $\int_0^1 t^5 dt$ donne, avec 1, 2 et 4 points de subdivision les valeurs

$$A_{0,0} = \frac{1}{2}, \quad A_{1,0} = \frac{17}{64} \quad \text{et} \quad A_{2,0} = \frac{197}{1024}$$

On peut alors commencer le remplissage du tableau comme au paragraphe précédent et on obtient la valeur exacte $A_{2,2} = \frac{1}{6}$. Il est d'ailleurs normal d'arriver à un tel résultat, puisque si f est un polynôme, $T_n(f, a, b)$ est un polynôme en $\frac{1}{n}$ et une itération suffisante du procédé de Richardson amène à une fonction constante.

$$\begin{array}{ccc} \frac{1}{2} & \frac{3}{16} & \frac{1}{6} \\ \frac{17}{64} & \frac{43}{256} & \\ \frac{197}{1024} & & \end{array}$$

EXERCICE 10-4.2 Montrer que, avec les notations précédentes, $A_{m,1}$ est l'approximation de Simpson avec 2^{m-1} points de subdivision.

EXERCICE 10-4.3 Programmer en Maple la formule des trapèzes et l'accélération de Romberg. Tester sur des exemples simples.

10-5 Exercices

EXERCICE 10-5.1 Soit f_n définie par $f_n(t) = \dfrac{n}{n^2t^2+1}$ et g continue de $[a, b]$ dans $\mathbb{R}$. Déterminer

$$\lim_{n\to+\infty} \frac{1}{\pi} \int_a^b f_n(t)g(t)dt$$

EXERCICE 10-5.2 On pose $a_n = \int_0^1 (1 - t + t^2)^n dt$. Montrer que la série $\sum(-1)^n a_n$ converge. Trouver une relation du type $a_n = f(n)a_{n-1} + g(n)$, où f et g sont des fonctions rationnelles. Montrer que $a_n \sim \frac{2}{n}$.

EXERCICE 10-5.3 Soit $n \in \mathbb{N}^*$ et $f \in C^0([a, b], \mathbb{R})$ telle que

$$\forall k \in \mathbb{N} \quad k \leqslant n \Longrightarrow \int_a^b t^k f(t)dt = 0$$

Montrer que f admet au moins $n + 1$ zéros sur $]a, b[$.

EXERCICE 10-5.4 Soient u et v les racines de l'équation $x^2 - x + \frac{1}{10} = 0$. Montrer que pour tout polynôme $P \in \mathbb{R}_5[X]$ on a $\int_0^1 P(t)dt = \frac{1}{18}\left[5P(u) + 8P(\frac{1}{2}) + 5P(v)\right]$.

EXERCICE 10-5.5 Soit f continue strictement positive sur $[0,1]$. Déterminer

$$\lim_{n\to+\infty}\left(\int_0^1 (f(x))^{\frac{1}{n}}dx\right)^n$$

EXERCICE 10-5.6 Soit $f \in \mathcal{C}^1([0,1],\mathbb{R})$. Donner un développement limité d'ordre 2 de $a_n = \int_0^1 x^n f(x)dx$.

EXERCICE 10-5.7 Soit f de classe $\mathcal{C}^1$ de $[0,+\infty[$ dans $\mathbb{R}$. On suppose f' croissante. Montrer que, pour tout $n \geqslant 2$:

$$0 \leqslant \frac{1}{2}f(1) + f(2) + \cdots + f(n-1) + \frac{1}{2}f(n) - \int_1^n f(t)dt \leqslant \frac{1}{8}(f'(n) - f'(1)).$$

EXERCICE 10-5.8 On pose $f(x) = \int_x^{x^2} \frac{dt}{\ln t}$ pour $x \in]0,1[\cup]1,+\infty[$.

1. Dérivée de f, sens de variation.
2. Montrer que $\lim_{x\to1} f(x) = \ln 2$.
3. Montrer que $f(x) \sim \frac{x^2}{2\ln x}$ quand $x \to +\infty$.
4. Trouver un équivalent de $f(x)$ pour $x \to 0$.

EXERCICE 10-5.9 Déterminer les triplets $(a,b,c) \in \mathbb{R}^3$ tels que $\int \frac{ax^2+bx+c}{(x^2+1)^2}dx$ soit une fonction rationnelle.

EXERCICE 10-5.10 Calculer l'intégrale $\int \frac{dt}{\sqrt{1+t^2}+\sqrt{1-t^2}}$.

EXERCICE 10-5.11 Etudier $f(x,y) = \int_0^\pi \ln(x + y\cos t)dt$ à l'aide d'une dérivation par rapport à x.

EXERCICE 10-5.12 Montrer que $u \mapsto \frac{\arctan(u)}{u}$ est C^∞ sur $\mathbb{R}$. En déduire le calcul de

$$\int_0^{\frac{\pi}{2}} \frac{\arctan(x\sin t)}{\sin t}dt$$

EXERCICE 10-5.13 Trouver un équivalent de $\int_0^\pi (\sin t)t^x dt$ lorsque x tend vers $+\infty$

EXERCICE 10-5.14 Soit f continue $[a,b] \to \mathbb{R}^{*+}$. Montrer qu'on définit une suite $(a_k)_{0\leqslant k\leqslant n}$ par $a_0 = a$ et $\int_{a_k}^{a_{k+1}} f(t)dt = \frac{1}{n}\int_a^b f(t)dt$. Etudier $\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(a_k)$.

EXERCICE 10-5.15 Soit f une application de classe C^1 sur $[a,b]$, à valeurs réelles. Déterminer la limite de la suite (nu_n) où $u_n = \int_a^b f(t)dt - \frac{b-a}{n}\sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right)$

EXERCICE 10-5.16 Soit $P \in \mathbb{R}[X]$ avec $P(X) = \sum_{n=0}^{n} a_k X^{2k}$. Montrer que

$$\int_0^1 P(t)dt = -\frac{i}{2}\int_0^{\pi} P(e^{i\theta})e^{i\theta}d\theta$$

Montrer que

$$\int_0^1 P^2(x)dx \leqslant \frac{\pi}{2}\sum_0^n |a_k|^2$$

Soit $f : [0,1] \to \mathbb{R}$ continue. On pose $I_n = \int_0^1 t^{2n} f(t)dt$. Montrer que la série $\sum I_n^2$ converge et que

$$\sum_{n=0}^{+\infty} I_n^2 \leqslant \frac{\pi}{2}\int_0^1 f^2(t)dt$$

EXERCICE 10-5.17 Calculer la primitive de $x \mapsto \frac{1}{3+\cos x}$ qui s'annule en 0.

Chapitre 11

Suites et séries de fonctions

11-1 Rappels : convergence simple et uniforme

Nous avons déjà introduit les notions de convergence simple et uniforme d'une suite de fonctions. Pour l'essentiel, les fonctions considérées dans ce chapitre seront des fonctions d'une variable réelle ou complexe à valeurs dans un espace normé (complet, le plus souvent $\mathbb{R}$ ou $\mathbb{C}$). Néanmoins, comme les résultats de ce chapitre pourront notamment s'appliquer à des fonctions de plusieurs variables réelles, nous nous placerons pour les définitions dans un cadre plus vaste : fonctions définies sur une partie d'un espace normé, à valeurs dans un espace normé (complet).

11-1.1 Convergence simple

Si X est une partie d'un espace normé $\mathbb{E}$, une suite d'applications $f_n : X \to \mathbb{F}$ (où $\mathbb{F}$ désigne un autre espace normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) converge simplement (sur X) vers une fonction $f : X \to \mathbb{F}$ ssi

$$\forall x \in X \quad \lim_{n \to +\infty} f_n(x) = f(x)$$

propriété que nous notons en abrégé

$$f_n \overset{\text{C.V.S}}{\underset{X}{\to}} f$$

Si cette notion est compatible avec les opérations de combinaisons linéaires que l'on peut considérer sur les fonctions (si $f_n \overset{C.V.S}{\underset{X}{\to}} f$ et $g_n \overset{C.V.S}{\underset{X}{\to}} g$ alors pour $\alpha \in \mathbb{K}$ on a $f_n + \alpha g_n \overset{C.V.S}{\underset{X}{\to}} f + \alpha g$) ainsi qu'avec d'autres opérations usuelles (produit dans le cas de fonctions réelles ou complexes, quotient dans la même situation si les dénominateurs ne s'annulent pas), les propriétés des fonctions f_n ne se conservent pas en général par passage à la limite simple :

- **Caractère borné :** si toutes les fonctions f_n sont bornées sur X et $f_n \overset{C.V.S}{\underset{X}{\to}} f$, la fonction f n'est pas bornée en général. Prendre par exemple

$$f_n(x) = \frac{1}{x + \frac{1}{n}} \text{ sur } X = \mathbb{R}^{*+}$$

- **Continuité :** si $f_n \overset{C.V.S}{\underset{X}{\to}} f$ et si toutes les f_n sont continues en un point $x_0 \in X$, il n'en est pas de même en général pour f. Prendre par exemple

$$f_n(x) = e^{-nx} \text{ sur } X = \mathbb{R}^+ \text{ avec } x_0 = 0$$

- **Intégration :** si $X = [a, b]$ est un segment, si les f_n sont continues par morceaux et convergent simplement sur $[a, b]$ vers une fonction f continue par morceaux[1], lorsque la limite des intégrales des f_n existe, on peut avoir

$$\lim_{n \to +\infty} \int_a^b f_n(t)\, dt \neq \int_a^b f(t)\, dt$$

Prendre par exemple

$$f_n = n.1_{]0, \frac{1}{n}[} \text{ sur } X = [0, 1]$$

(en prenant $g_n = (-1)^n f_n$, on aurait un exemple de suite convergeant simplement sur $[0, 1]$ vers la fonction nulle, mais pour laquelle la suite des intégrales $\int_0^1 g_n(t)\, dt$ n'a pas de limite).

- **Dérivabilité :** si $X = [a, b]$ est un segment et $f_n \overset{C.V.S}{\underset{[a,b]}{\to}} f$ avec f_n dérivable sur X, il n'en est pas de même en général pour f. Prendre

$$f_n(x) = \sqrt{x^2 + \frac{1}{n}} \text{ sur } [-1, 1]$$

[1]Cette propriété n'est pas assurée par la convergence simple : prendre par exemple $X = [0, 1]$ et $f_n(x) = 0$ sauf si x est un rationnel de $[0, 1]$ pouvant s'écrire $\frac{p}{q}$ avec $q \leqslant n$ auquel cas $f_n(x) = 1$. Il est facile de voir que f_n est continue par morceaux (c'est une fonction en escalier) mais

$$f_n \overset{C.V.S}{\to} 1_{\mathbb{Q} \cap [0,1]}$$

fonction caractéristique de l'ensemble des rationnels de $[0, 1]$, qui n'est pas continue par morceaux

Si on suppose de plus f dérivable, on n'aura pas nécessairement convergence simple de la suite f_n' vers f'. Prendre comme exemple

$$f_n(x) = \frac{\sin nx}{n} \text{ sur } X = \mathbb{R}$$

Pour obtenir des propriétés de régularité de la limite f, on fera souvent intervenir la notion plus fine de convergence uniforme.

11-1.2 Convergence uniforme

11-1.2.1 Définition

Nous avons défini (*cf.* théorème **6-2**.16) l'espace $\mathcal{B}(X, \mathbb{F})$ des fonctions bornées de X dans $\mathbb{F}$ et vu que, sur cet espace

$$\|f\|_\infty = \sup_{x \in X} \|f(x)\|$$

définit une norme, appelée **norme de la convergence uniforme sur** X. Lorsqu'on travaille en restriction à une partie $A \subset X$, on utilise aussi la norme de la convergence uniforme sur A

$$\|f\|_{\infty,A} = \sup_{x \in A} \|f(x)\|$$

(il ne s'agit pas d'une norme sur $\mathcal{B}(X, \mathbb{F})$, mais d'une semi-norme).

Une suite de fonctions $f_n : X \to \mathbf{F}$ converge uniformément sur X vers une fonction f si et seulement si

$$\begin{cases} \exists n_0 \in \mathbb{N} \quad \forall n \geqslant n_0 \quad (f - f_n) \in \mathcal{B}(X, \mathbf{E}) \\ \lim_{n \to +\infty} \|f - f_n\|_\infty = 0 \end{cases}$$

On notera en abrégé

$$f_n \overset{\text{C.V.U}}{\underset{X}{\to}} f \quad \text{ou} \quad f_n \overset{\text{C.V.U}}{\to} f$$

s'il ne risque pas d'y avoir ambiguité sur le domaine d'étude : **lorsqu'on étudie une convergence uniforme, il faut être très précis : quelle est la suite de fonction étudiée ?** (cette suite doit être explicitée). **Sur quel domaine y a-t-il convergence uniforme ?** (toute affirmation relative à une convergence uniforme doit être justifiée).

Pour que $f_n \overset{\text{C.V.U}}{\underset{X}{\to}} f$, il est nécessaire et suffisant qu'il existe une suite réelle positive $(\alpha_n)_{n \in \mathbb{N}}$ telle que

$$\lim_{n \to +\infty} \alpha_n = 0 \text{ et } \forall x \in X \quad \|f(x) - f_n(x)\| \leqslant \alpha_n$$

le terme général de cette suite dépend évidemment de n, mais pas de x.

De même, pour montrer qu'il n'y a pas convergence uniforme sur X, il suffit d'exhiber une suite (x_n) de points de X telle que la suite $(\|f(x_n) - f_n(x_n)\|)_{n \in \mathbb{N}}$

ne tende pas vers 0. Lorsque les fonctions f et f_n sont des fonctions réelles d'une variable réelle, c'est parfois l'étude des variations de la fonction $f - f_n$ qui permet de déterminer un tel x_n.

Rappelons enfin que la convergence uniforme d'une suite de fonctions sur une **famille finie de parties** de X entraîne la convergence uniforme sur **la réunion de ces parties**, mais qu'il n'en est pas de même pour une famille infinie.

EXERCICE 11-1.1 Montrer que la suite de fonctions $f_n : \mathbb{R}^+ \to \mathbb{R}$ définie par

$$f_n(x) = \frac{x}{x+n}$$

converge simplement sur $\mathbb{R}^+$ vers la fonction nulle, uniformément sur tout segment $[0, a]$ avec $a > 0$, mais pas uniformément sur $[0, +\infty[= \bigcup_{a>0} [0, a]$.

11-1.2.2 Critère de Cauchy uniforme

Nous avons vu (théorème **6-5**.10) que, lorsque l'espace d'arrivée $\mathbb{F}$ est complet, l'espace $\mathcal{B}(X, \mathbb{F})$ est complet pour $\| \|_\infty$. Ceci nous donne un critère de convergence uniforme (très souvent utilisé dans le cadre des séries de fonctions, de manière plus ou moins explicite) utile lorsqu'on ne connaît pas *a priori* la fonction limite f :

PROPOSITION 11-1.2 Si l'espace d'arrivée $\mathbb{F}$ est complet, une suite de fonctions $f_n : X \to \mathbb{F}$ converge uniformément sur X si et seulement si

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n, p \geqslant N \quad \forall x \in X \quad \|f_n(x) - f_p(x)\| \leqslant \varepsilon$$

On notera bien la place des quantificateurs : N ne dépend que de ε. Cette proposition équivaut évidemment à

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n, p \geqslant N \quad (f_n - f_p) \in \mathcal{B}(X, \mathbb{F}) \text{ et } \|f_n - f_p\|_\infty \leqslant \varepsilon$$

Démonstration : Si la suite converge uniformément vers une fonction f sur X et si $\varepsilon > 0$ est fixé, on peut trouver un rang N tel que

$$\forall n \geqslant N \quad (f - f_n) \in \mathcal{B}(X, \mathbb{F}) \text{ et } \|f - f_n\|_\infty \leqslant \frac{\varepsilon}{2}$$

On a alors, pour n et $p \geqslant N$

$$(f_n - f_p) = (f - f_p) - (f - f_n) \in \mathcal{B}(X, \mathbb{F})$$
$$\text{et } \|f_n - f_p\|_\infty \leqslant \|f_n - f\|_\infty + \|f - f_p\|_\infty \leqslant \varepsilon$$

Réciproquement, si la suite (f_n) vérifie les hypothèses de la proposition, en prenant $\varepsilon = 1$ on trouve un entier N_1 tel que

$$n \geqslant N_1 \Rightarrow (f_n - f_{N_1}) \in \mathcal{B}(X, \mathbb{F})$$

La suite $(g_n)_{n \geqslant N_1}$ avec $g_n = f_n - f_{N_1}$ est alors de Cauchy dans $(\mathcal{B}(X, \mathbb{F}), \| \|_\infty)$, puisque

$$\forall n, p \geqslant N_1 \quad \|g_n - g_p\| = \|f_n - f_p\|$$

Comme on suppose $\mathbb{F}$ complet, il en est de même de $\mathcal{B}(X,\mathbf{F})$. La suite (g_n) est donc uniformément convergente sur X vers une fonction g. Il est alors facile de voir que la suite (f_n) converge uniformément vers $f = g + f_{N_1}$. ■

EXERCICE 11-1.3 Lorsque l'espace $\mathbf{F}$ n'est pas complet, montrer que la convergence simple et le critère de Cauchy uniforme entraînent la convergence uniforme.

11-1.2.3 Convergence uniforme et opérations algébriques

Lorsqu'on étudie une convergence uniforme sur X, on se ramène toujours, par translation, à travailler avec des fonctions de $\mathcal{B}(X,\mathbb{F})$. Comme la convergence uniforme est alors convergence pour une norme, elle est compatible avec les opérations de cet espace vectoriel :

PROPOSITION 11-1.4 Si $f_n \overset{\text{C.V.U}}{\underset{X}{\to}} f$ et $g_n \overset{\text{C.V.U}}{\underset{X}{\to}} g$, alors pour $\alpha \in \mathbb{K}$

$$f_n + \alpha g_n \overset{\text{C.V.U}}{\underset{X}{\to}} f + \alpha g$$

Nous limitons l'énoncé qui suit au cas (le plus important en pratique) de fonctions à valeurs réelles ou complexes. Il y a des extensions évidentes aux cas de fonctions à valeurs dans une algèbre normée, du produit d'une fonction scalaire par une fonction vectorielle ou plus généralement de l'utilisation d'une application bilinéaire continue entre espaces normés[2] :

PROPOSITION 11-1.5 Soient (f_n) et (g_n) deux suites de fonctions de X dans $\mathbb{R}$ ou $\mathbb{C}$ convergeant uniformément sur X vers des fonctions f et g. Si à partir d'un certain rang les fonctions f_n et g_n sont bornées (ou, ce qui revient au même, si f et g sont bornées) alors

$$f_n g_n \overset{\text{C.V.U}}{\underset{X}{\to}} fg$$

Le résultat ne persiste pas en général si l'une des fonctions f ou g n'est pas bornée.

Démonstration : Avec les hypothèses faites sur (f_n) et (g_n), tout se passe dans l'espace normé $\mathcal{B}(X,\mathbf{F})$. Sur cet espace muni de la norme de la convergence uniforme, le produit est une application bilinéaire et continue puisque

$$\forall \varphi, \psi \in \mathcal{B}(X,\mathbb{F}) \quad \|\varphi\psi\|_\infty \leqslant \|\varphi\|_\infty \|\psi\|_\infty$$

Donc si $f_n \to f$ et $g_n \to g$ pour la norme $\| \|_\infty$ alors $f_n g_n \to fg$ pour cette norme. Dans le cas $X = \mathbb{R}$, l'exemple

$$f_n(x) = x + \frac{1}{n} \text{ et } g_n(x) = \frac{1}{n}$$

montre que le résultat n'est plus vrai en général lorsqu'une des fonctions f ou g n'est pas bornée. ■

[2] C'est en effet la généralisation de la notion de produit.

COROLLAIRE 11-1.6 On retiendra en particulier de la démonstration précédente que, si $f_n \overset{\text{C.V.U}}{\underset{X}{\to}} f$ et si g est bornée, alors $f_n g \overset{\text{C.V.U}}{\underset{X}{\to}} fg$, à cause de la majoration

$$\|f_n g - fg\|_\infty \leqslant \|f_n - f\|_\infty \|g\|_\infty$$

La justification "*g* **est bornée**" doit apparaître explicitement dans le raisonnement : méditer le contre exemple $X = \mathbb{R}$, $g(x) = x$ et $f_n(x) = \dfrac{\sin nx}{n}$.

11-1.3 Propriétés de la convergence uniforme

11-1.3.1 Continuité, interversion de limites

Rappelons les résultats obtenus aux sections **7-3.2** et **7-3.3** :

- Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions $X \to \mathbb{F}$ qui converge uniformément sur X vers une fonction f. Si toutes les f_n sont continues en un point $a \in X$, alors f est continue en a.

- Une limite uniforme d'une suite de fonctions continues sur X est continue sur X.

- Soit $\mathbb{F}$ un espace normé complet, A une partie d'un autre espace normé et $f_n : A \to \mathbb{F}$ une suite de fonctions convergeant uniformément sur A vers une fonction f. Soit a un point de $\overline{A} - A$, adhérent à A. On suppose que chacune des fonctions f_n possède une limite $l_n \in \mathbb{F}$ en a. Alors la suite $(l_n)_{n\in\mathbb{N}}$ est convergente dans $\mathbb{F}$, la fonction f possède une limite en a et

$$\lim_{x\to a} f(x) = \lim_{x\to a} \lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} l_n = \lim_{n\to+\infty} \lim_{x\to a} f_n(x)$$

Quelques commentaires sur ces résultats :

- Lorsque la conclusion d'un de ces énoncés est en défaut alors qu'il y a convergence simple et que toutes les autres hypothèses sont vérifiées, on est assuré qu'il n'y a pas convergence uniforme.

- Dans les énoncés précédents, les hypothèses sont des conditions suffisantes et pas des conditions nécessaires pour pouvoir conclure. Par exemple, la suite de fonctions continues de $\mathbb{R}$ dans lui-même définies par

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n$$

converge simplement sur $\mathbb{R}$ vers la fonction continue $f(x) = e^x$, mais il n'y a certainement pas convergence uniforme (pourquoi ?)[3]. Il faut donc être précis dans ses affirmations : après avoir vérifié la continuité de chaque f_n en un point a, pour être assuré de la continuité de la limite simple f en a, on dira : "il **suffit** de vérifier la convergence uniforme sur un voisinage de a" et non pas "**il faut** ..."[4].

[3] Il y a cependant (exercice !) convergence uniforme locale (c'est à dire au voisinage de tout réel x_0) ce qui est suffisant pour assurer la continuité de la limite en tout point x_0.

[4] Même si c'est souvent ce qu'il "faut faire" pour se sortir d'une situation délicate : le langage courant est moins précis que celui des énoncés mathématiques.

EXERCICE 11-1.7 Montrer que la limite simple de la suite de fonctions $f_n : \mathbb{R}^+ \to \mathbb{R}$

$$f_n(x) = n^2 x e^{-nx}$$

est continue en 0 alors qu'il n'y a convergence uniforme sur aucun voisinage de 0.

11-1.3.2 Intégration sur un segment

Rappelons le résultat du théorème **10-1**.5 (l'espace **F** est supposé complet, pour pouvoir construire l'intégrale d'une fonction continue par morceaux) :

Si $f_n : [a, b] \to \mathbb{F}$ est une suite de fonctions continues par morceaux qui converge uniformément sur $[a, b]$ vers une fonction f continue par morceaux, alors

$$\lim_{n\to+\infty} \int_a^b f_n(t)\, dt = \int_a^b f(t)\, dt$$

ce qui revient à intervertir les symboles d'intégration et de passage à la limite. On peut d'ailleurs être plus précis en disant que la suite f_n converge vers f en moyenne sur $[a, b]$ (ce qui entraîne la convergence des intégrales) puisque

$$\|f - f_n\|_1 = \int_a^b |f(t) - f_n(t)|\, dt \leqslant (b - a)\, \|f - f_n\|_\infty$$

EXERCICE 11-1.8 Si x est un réel positif ou nul, montrer que

$$\lim_{n\to+\infty} \int_0^1 t^x \left(1 - \frac{t}{n}\right)^n dt = \int_0^1 t^x e^{-t}\, dt$$

(Remarque : nous verrons, dans un chapitre ultérieur sur l'intégration sur un intervalle quelconque, des théorèmes qui évitent souvent le recours à l'étude de la convergence uniforme).

EXERCICE 11-1.9 La suite de fonctions définies par

$$f_n(x) = \frac{2^n x}{1 + n 2^n x^2}$$

converge-t-elle uniformément sur $[0, 1]$?

11-1.3.3 Dérivation et convergence uniforme

Comme le montre l'exemple donné en début de ce chapitre, une limite uniforme sur un intervalle d'une suite de fonctions dérivables n'est pas nécessairement dérivable : la suite de fonctions définies sur $[-1, 1]$ par

$$f_n(x) = \sqrt{x^2 + \frac{1}{n}}$$

est une suite de fonctions de classe C^1 qui converge uniformément sur $[-1, 1]$ (exercice) vers f définie par $f(x) = |x|$ qui n'est pas dérivable en 0.

Pour établir la dérivabilité d'une limite d'une suite de fonctions, on dispose du théorème suivant, où l'on notera que **l'hypothèse de convergence uniforme**

porte sur la suite des dérivées. Ici encore, ce théorème donne une condition **suffisante** pour pouvoir permuter les opérations de passage à la limite et de dérivation.

Nous nous limiterons aux suites de fonctions de classe $\mathcal{C}^1$, ce qui permet d'écrire chaque fonction comme intégrale de sa dérivée, et ramène ce résultat au théorème sur l'intégration des suites uniformément convergentes. Bien entendu, l'espace $\mathbb{F}$ est supposé complet.

THÉORÈME 11-1.10 Soit I un intervalle de $\mathbb{R}$, et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de classe $\mathcal{C}^1$ de I dans $\mathbb{F}$. On suppose

1) Il existe un point $a \in I$ tel que la suite $(f_n(a))_{n\in\mathbb{N}}$ soit convergente dans $\mathbb{F}$

2) La suite des dérivées $(f_n')_{n\in\mathbb{N}}$ converge uniformément vers une fonction g sur tout segment $\subset I$

Alors la suite (f_n) converge simplement sur I vers une fonction f, avec convergence uniforme sur tout segment inclus dans I. La fonction f est de classe $\mathcal{C}^1$ sur I, avec $f' = g$, ce qui peut s'écrire en abrégé

$$\left(\lim_{n\to+\infty} f_n\right)' = \lim_{n\to+\infty} f_n'$$

Démonstration : La convergence uniforme locale de la suite de fonctions continues $(f_n')_{n\in\mathbb{N}}$ assure la continuité de g. Si l est la limite de la suite $(f_n(a))_{n\in\mathbb{N}}$, posons

$$\forall x \in I \quad f(x) = l + \int_a^x g(t)\, dt$$

Cette fonction est de classe $\mathcal{C}^1$ et vérifie $f' = g$. Comme on a également

$$\forall x \in I \quad f_n(x) = f_n(a) + \int_a^x f_n'(t)\, dt$$

si K est un segment inclus dans I nous aurons

$$\forall x \in K \quad \|f_n(x) - f(x)\| \leqslant \|f_n(a) - l\| + \left\| \int_a^x (f_n'(t) - g(t))\, dt \right\|$$

et si K_1 est le plus petit segment contenant a et K, sa longueur étant notée $L(K_1)$

$$\forall x \in K \quad \|f_n(x) - f(x)\| \leqslant \|f_n(a) - l\| + L(K_1)\, \|f_n' - g\|_{\infty, K_1}$$

Cette inégalité prouve la convergence uniforme de la suite (f_n) vers f sur K. ■

REMARQUE 11-1.11 On comprend bien l'hypothèse relative à l'existence du point a : si on ne conserve que l'hypothèse sur les dérivées, on peut ajouter une suite arbitraire $(a_n)_{n\in\mathbb{N}}$ de $\mathbb{F}^{\mathbb{N}}$ à la suite $(f_n)_{n\in\mathbb{N}}$ ce qui laisse peu de chances de voir converger simplement cette suite de fonctions.

REMARQUE 11-1.12 Dans la pratique, le théorème qui précède est utilisé de manière un peu affaiblie. On a une suite (f_n) de fonctions de classe C^1 de I dans $\mathbf{F}$ dont on vérifie d'abord la convergence (en général simple) sur I vers une fonction f. On se pose la question de la régularité de f : f est-elle dérivable, comment évaluer f' ? Pour que f soit dérivable (et même de classe C^1) sur I, **il suffit que la suite des dérivées $(f'_n)_{n\in\mathbb{N}}$ converge uniformément (localement) sur I.** Et on a alors accès à f' par

$$f' = \lim_{n\to+\infty} f'_n$$

Le théorème qui précède est rarement utilisé pour justifier l'existence de la fonction f, limite simple de la suite (f_n).

COROLLAIRE 11-1.13 Soit $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions de classe C^p ($p \geqslant 1$) de I dans $\mathbb{F}$ telle que, pour $0 \leqslant k \leqslant p$, les suites $\left(f_n^{(k)}\right)_{n\in\mathbb{N}}$ convergent simplement vers une fonction f_k, la convergence étant uniforme sur tout segment $\subset I$ pour la suite des dérivées d'ordre p. La fonction f_0 est alors de classe C^p sur I et, pour tout k vérifiant $1 \leqslant k \leqslant p$,

$$f_0^{(k)} = f_k \quad \textbf{soit} \quad \left(\lim_{n\to+\infty} f_n\right)^{(k)} = \lim_{n\to+\infty} f_n^{(k)}$$

Démonstration : par récurrence sur p. ∎

11-2 Cas des séries de fonctions

Nous avons vu au chapitre consacré sur les séries dans un espace vectoriel normé qu'il n'y a pas de différence fondamentale entre l'étude d'une suite et l'étude d'une série, la différence se situant essentiellement au niveau du vocabulaire : étudier la série de terme général $u_n \in \mathbb{F}$, c'est étudier la suite $(S_n)_{n\in\mathbb{N}}$ des sommes partielles

$$S_n = \sum_{k=0}^{n} u_k$$

et de même, l'étude de la suite de terme général u_n peut se ramener à l'étude de la série télescopique de terme général $v_n = u_n - u_{n-1}$ ($n \geqslant 1$) avec $v_0 = u_0$.

Nous allons donc traduire, avec le vocabulaire des séries, les résultats de la section précédente.

11-2.1 Convergence uniforme d'une série

11-2.1.1 Définition

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions : $X \to \mathbb{F}$. Etudier la série de fonctions $\sum u_n(x)$ c'est d'abord déterminer l'ensemble des $x \in X$ tels que cette série (de vecteurs de $\mathbb{F}$) converge :

$$Y = \left\{x \in X \mid \sum u_n(x) \ \text{CV}\right\}$$

On peut alors définir une fonction S (somme de la série)

$$S : Y \to \mathbf{F} \quad S(x) = \sum_{n=0}^{+\infty} u_n(x)$$

et on dit alors que la série de fonctions de terme général u_n converge simplement sur Y vers S. Cela signifie simplement que, pour tout $x \in Y$, la suite des sommes partielles

$$\left(S_n(x) = \sum_{k=0}^{n} u_k(x)\right)_{n\in\mathbb{N}}$$

converge vers $S(x)$, ce qui revient à affirmer la convergence simple de la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ vers S sur Y :

Etudier la convergence simple d'une série de fonctions à valeurs dans $\mathbb{F}$, c'est étudier une série de vecteurs de $\mathbb{F}$ dépendant d'un paramètre x. Pour ce faire, on dispose de tous les théorèmes, critères etc... qui ont été vus dans le chapitre sur les séries. En général, cette étude est faite avant de se poser le problème de la convergence uniforme :

DÉFINITION 11-2.1 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de X dans $\mathbf{F}$. On dit que la série de fonctions $\sum u_n$ (ou, par abus d'écriture $\sum u_n(x)$) converge uniformément sur X si et seulement si $\forall x \in \mathbb{X}$ $\sum u_n(x)$ converge et la suite des sommes partielles $S_n : x \mapsto S_n(x) = \sum_{k=0}^{n} u_k(x)$ converge uniformément sur X.*

Comme la limite des sommes partielles est, par définition,

$$S(x) = \sum_{n=0}^{+\infty} u_n(x)$$

on a

$$S(x) - S_n(x) = \sum_{k=n+1}^{+\infty} u_k(x) = R_n(x) \quad \text{(reste d'ordre } n \text{ de la série)}$$

et donc **la série $\sum u_n(x)$ converge uniformément sur X si et seulement si la suite des restes converge uniformément vers la fonction nulle :**

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \forall x \in X \quad \|R_n(x)\| \leqslant \varepsilon$$

(il est bien évident qu'on ne peut parler de reste de la série qu'après avoir prouvé la convergence).

Prouver la convergence uniforme d'une série, c'est donc quasiment toujours majorer uniformément les restes $(\|R_n(x)\|)_{n\in\mathbb{N}}$ par une suite $(\alpha_n)_{n\in\mathbb{N}}$ tendant vers 0 et **ne dépendant pas de** x (cette majoration est parfois implicite, par exemple

lorsqu'on applique la convergence normale que nous verrons dans la section qui suit, mais il importe d'avoir toujours cette idée simple à l'esprit).

Lorsque l'espace $\mathbb{F}$ est complet, le critère de Cauchy uniforme est aussi un moyen de prouver à la fois la convergence simple et la convergence uniforme d'une série :

PROPOSITION 11-2.2 Si $\mathbb{F}$ est complet, une série de fonctions $\sum u_n$, avec $u_n : X \to \mathbb{F}$ converge uniformément sur X si et seulement si

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \forall p \geqslant 0 \quad \forall x \in X$$
$$\|u_n(x) + u_{n+1}(x) + \cdots + u_{n+p}(x)\| \leqslant \varepsilon$$

Même lorsque l'espace n'est pas complet, la condition qui précède est nécessaire pour que la série $\sum u_n$ converge uniformément sur X. En particulier, en appliquant ce résultat avec $p = 0$, on obtient :

PROPOSITION 11-2.3 Si une série de fonctions converge uniformément, son terme général converge uniformément vers la fonction nulle.

Il n'y a évidemment pas de réciproque, puisque le fait que le terme général tende vers 0 n'assure même pas la convergence de la série.

EXEMPLE 11-2.4 La série géométrique $\sum x^n$ converge simplement sur $]-1, 1[$, mais pas uniformément.

11-2.1.2 Cas particulier de la convergence normale

On suppose ici l'espace $\mathbb{F}$ complet.

DÉFINITION 11-2.5 *Si $(u_n)_{n \in \mathbb{N}}$ est une suite de fonctions de X dans $\mathbb{F}$, la série de fonctions $\sum u_n$ converge normalement sur X si et seulement s'il existe un rang n_0 tel que*

$$n \geqslant n_0 \Rightarrow u_n \in \mathcal{B}(X, \mathbb{F}) \quad \textit{et} \quad \sum_{n \geqslant n_0} \|u_n\|_\infty \ \textit{converge}$$

(l'hypothèse $u_n \in \mathcal{B}(X, \mathbb{F})$ est là uniquement pour donner un sens à $\|u_n\|_\infty$).

REMARQUE 11-2.6 Dans la pratique, on ne cherche pas à expliciter $\|u_n\|_\infty$: il suffit de trouver une suite $(\alpha_n)_{n \geqslant n_0}$ telle que

$$\sum_{n \geqslant n_0} \alpha_n \text{ converge} \quad \text{et} \quad \forall n \geqslant n_0 \quad \forall x \in X \quad \|u_n(x)\| \leqslant \alpha_n$$

En effet, si la série converge normalement, on peut prendre $\alpha_n = \|u_n\|_\infty$. Réciproquement, si une telle suite $(\alpha_n)_{n \geqslant n_0}$ existe, on aura $\|u_n\|_\infty \leqslant \alpha_n$ pour $n \geqslant n_0$, ce qui assure, par majoration, la convergence de la série (à termes positifs !) $\sum_{n \geqslant n_0} \|u_n\|_\infty$

On majore donc, INDEPENDAMMENT DE x, la norme du terme général de la série de fonctions par le terme général (positif) d'une série convergente.

L'intérêt de cette notion est d'offrir une condition suffisante simple de convergence uniforme :

THÉORÈME 11-2.7 Si $\mathbb{F}$ est complet, toute série de fonctions de X dans $\mathbb{F}$ qui converge normalement sur X est uniformément convergente sur X.

Démonstration : Si $\sum_{n \geqslant n_0} \|u_n\|_\infty$ est convergente, on a d'abord

$$\forall n \geqslant n_0 \quad \forall x \in X \quad \|u_n(x)\| \leqslant \|u_n\|_\infty$$

majoration qui assure la convergence absolue (donc la convergence puisque $\mathbb{F}$ est complet) de la série $\sum u_n(x)$, pour tout $x \in X$. De plus, pour $n \geqslant n_0$,

$$\|R_n(x)\| = \left\| \sum_{k=n+1}^{+\infty} u_k(x) \right\| \leqslant \sum_{k=n+1}^{+\infty} \|u_k(x)\| \leqslant \sum_{k=n+1}^{+\infty} \|u_k\|_\infty$$

majoration du reste de la série par une suite tendant vers 0 indépendante de x : la convergence uniforme de la série est prouvée[5]. ■

EXEMPLE 11-2.8 La série de fonctions de la variable réelle x

$$\sum_{n=1}^{+\infty} \frac{1}{n^2 + x^2}$$

converge normalement, et est donc uniformément convergente sur $\mathbb{R}$.

Attention ! La convergence normale sur X est une condition suffisante de convergence uniforme sur X, qui entraîne en particulier la convergence absolue de la série $\sum u_n(x)$ pour tout $x \in X$. Cette remarque montre que **la convergence normale n'est pas une condition nécessaire de convergence uniforme,** puisqu'on peut facilement trouver des séries uniformément

[5]On aurait pu faire plus savant : l'hypothèse

$$\sum_{n=n_0}^{+\infty} \|u_n\|_\infty < +\infty$$

traduit, dans l'espace normé $(\mathcal{B}(X,\mathbb{F}), \|\,\|_\infty)$ la convergence **absolue** de la série $\sum_{n=n_0}^{+\infty} u_n$. L'espace $\mathcal{B}(X,\mathbb{F})$ étant complet pour la norme de la convergence uniforme, cette convergence absolue entraîne la convergence (pour $\|\,\|_\infty$) de la série $\sum_{n=n_0}^{+\infty} u_n$, ce qui revient exactement à la convergence uniforme de la série de fonctions $\sum_{n=n_0}^{+\infty} u_n(x)$.

convergentes qui ne sont pas absolument convergentes à x fixé. L'exemple le plus simple est celui de la série de fonctions de la variable réelle x

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n}$$

qui converge (par application du théorème des séries alternées) uniformément sur $\mathbb{R}$, puisque son terme général ne dépend pas de x ! Pour un exemple un peu moins évident, voir la section qui suit.

11-2.1.3 Utilisation du théorème des séries alternées

L'énoncé qui suit s'appliquera dans un certain nombre de situations concrètes. Pour éviter des erreurs, il sera préférable de refaire la démonstration au cas par cas :

PROPOSITION 11-2.9 Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de X dans $\mathbb{R}^+$. On suppose qu'il existe un rang p INDEPENDANT DE x tel que

$$\forall x \in X \quad (u_n(x))_{n\geqslant p} \text{ est décroissante}$$

Alors la série $\sum (-1)^n u_n(x)$ est uniformément convergente sur X si et seulement si la suite de fonctions $(u_n)_{n\in\mathbb{N}}$ converge uniformément sur X vers la fonction nulle.

Démonstration : Cette condition est évidemment nécessaire d'après la proposition **11-2.3**. Si réciproquement $u_n \overset{\text{C.V.U}}{\underset{X}{\rightarrow}} 0$, pour tout $x \in X$ la série $\sum (-1)^n u_n(x)$ converge par application du théorème des séries alternées, et

$$\forall x \in X \quad \forall n \geqslant p \quad |R_n(x)| \leqslant |u_{n+1}(x)| \leqslant \|u_{n+1}\|_\infty$$

ce qui donne $\|R_n\|_\infty \leqslant \|u_{n+1}\|_\infty$. La série converge uniformément. ■

EXEMPLE 11-2.10 La série de fonctions de la variable x

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{x^2+n}$$

converge uniformément, mais pas normalement sur $\mathbf{R}$.

Attention ! Si le rang N_x à partir duquel la valeur absolue $u_n(x)$ du terme général commence à décroître dépend de x, on obtiendra une majoration du type

$$\forall x \in X \quad \forall n \geqslant N_x \quad |R_n(x)| \leqslant |u_{n+1}(x)| \leqslant \|u_{n+1}\|_\infty$$

qui ne permet pas en général de prouver la convergence uniforme si la correspondance $x \mapsto N_x$ n'est pas bornée :

EXERCICE 11-2.11 Montrer que la convergence uniforme sur $\mathbb{R}$ de la série de fonctions

$$\sum_{n=1}^{+\infty} (-1)^{n-1} \frac{n}{n^2+x^2}$$

ne peut se prouver par la technique précédente. On peut prouver cette convergence uniforme par une majoration du reste en groupant les termes deux par deux. Il est préférable d'écrire

$$u_n(x) = (-1)^{n-1} \frac{n}{n^2+x^2} = \frac{(-1)^{n-1}}{n} + v_n(x)$$

et décomposer la série en deux séries uniformément convergentes, en appliquant la proposition qui suit.

11-2.1.4 Théorèmes sur les séries uniformément convergentes

Nous traduisons ici, avec le vocabulaire des séries, les résultats de la section **11-1.3**. Rappelons une fois encore que les hypothèses de ces théorèmes ne sont que des conditions suffisantes de leurs conclusions :

Stabilité par combinaisons linéaires.

PROPOSITION 11-2.12 L'ensemble des suites de fonctions

$$\left\{(u_n)_{n\in\mathbb{N}} \in \left[\mathbb{F}^X\right]^{\mathbb{N}} \mid \sum u_n \text{ converge uniformément sur } X\right\}$$

est un espace vectoriel. Lorsque l'espace $\mathbb{F}$ est complet, l'ensemble des suites pour lesquelles la série est normalement convergente en est un sous-espace vectoriel.

Continuité.

PROPOSITION 11-2.13 Si $(u_n)_{n\in\mathbb{N}}$ est une suite d'applications $X \to \mathbb{F}$ qui sont toutes continues en $a \in X$, si la série $\sum u_n$ converge simplement sur X et uniformément sur un voisinage de a, alors la somme S de la série

$$S(x) = \sum_{n=0}^{+\infty} u_n(x)$$

est continue en a.

Démonstration : Il suffit d'appliquer le théorème correspondant sur les suites à $(S_n)_{n\in\mathbb{N}}$ avec

$$S_n(x) = \sum_{k=0}^{n} u_k(x) \quad \blacksquare$$

COROLLAIRE 11-2.14 Si les u_n sont toutes continues sur X et si $\sum u_n$ converge uniformément sur X, alors S est continue sur X.

EXEMPLE 11-2.15 Les fonctions définies sur $\mathbb{R}$ par

$$f(x)=\sum_{n=1}^{+\infty}\frac{1}{n^2+x^2}\quad\text{et}\quad g(x)=\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{n}{n^2+x^2}$$

sont continues sur $\mathbb{R}$.

Interversion de lim et de $\sum$

PROPOSITION 11-2.16 On suppose que l'espace $\mathbb{F}$ est complet. Soient $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions $A\to\mathbb{F}$ telle que $\sum u_n$ converge simplement sur A, et $a\in\overline{A}-A$ un point adhérent à A. On suppose que la série de fonctions $\sum u_n$ converge uniformément sur un ensemble de la forme $V\cap A$, avec V voisinage de a. On suppose de plus que toutes les fonctions u_n possèdent une limite en a

$$l_n=\lim_{\substack{x\to a\\ x\in A}}u_n(x)$$

Alors la somme $S:A\to\mathbb{F}$ de la série possède une limite en a, la série de terme général l_n est convergente et

$$\lim_{\substack{x\to a\\ x\in A}}S(x)=\lim_{\substack{x\to a\\ x\in A}}\sum_{n=0}^{+\infty}u_n(x)=\sum_{n=0}^{+\infty}l_n$$

ce qui peut s'écrire aussi

$$\lim_{\substack{x\to a\\ x\in A}}\sum_{n=0}^{+\infty}u_n(x)=\sum_{n=0}^{+\infty}\lim_{\substack{x\to a\\ x\in A}}u_n(x)$$

(si A est une partie non bornée[6] de l'espace normé de départ, on a un résultat analogue avec les limites à l'infini).

EXERCICE 11-2.17 Montrer que

$$S(x)=\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n}$$

est une série qui converge normalement sur tout segment $[-a,a]$ avec $0<a<1$, et en déduire que S est définie et continue sur $]-1,1[$. Montrer que

$$\lim_{\substack{x\to 1\\ x<1}}S(x)=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n}$$

EXERCICE 11-2.18 Si $A\in\mathcal{M}_p(\mathbb{C})$, montrer que

$$\lim_{n\to+\infty}\left(I_p+\frac{A}{n}\right)^n=\exp(A)$$

[6] Pour nous, le plus souvent, un intervalle non borné de $\mathbb{R}$, ou $\mathbb{N}$.

Indication : on peut utiliser ici un argument de convergence uniforme. En utilisant la formule du binôme, écrire

$$\left(I_p + \frac{A}{n}\right)^n = \sum_{k=0}^{+\infty} u_k(n)$$

où $u_k(n) = \alpha_k(n) A^k$ est une suite nulle à partir du rang $n+1$ (cette somme ne fait intervenir qu'un nombre fini de termes, mais comme ce nombre croît indéfiniment avec n il est intéressant d'utiliser la notion de somme d'une série). Montrer que cette série de fonctions de la variable entière n converge uniformément sur $X = \mathbb{N}$. Conclure.

EXERCICE 11-2.19 Déterminer

$$\lim_{x\to+\infty} \sum_{n=1}^{+\infty} (-1)^n \frac{n}{n^2+x^2}$$

EXERCICE 11-2.20 En utilisant l'encadrement

$$\int_n^{n+1} \frac{dt}{t^2+x^2} \leqslant \frac{1}{n^2+x^2} \leqslant \int_{n-1}^{n} \frac{dt}{t^2+x^2}$$

déterminer

$$\lim_{x\to+\infty} \sum_{n=1}^{+\infty} \frac{x}{n^2+x^2}$$

Qu'aurait-on obtenu en permutant sans précaution les symboles de sommation et de passage à la limite ?

Intégration terme à terme.

PROPOSITION 11-2.21 Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'applications continues par morceaux de $[a,b]$ dans $\mathbb{F}$. On suppose que la série de fonction $\sum u_n$ est uniformément convergente sur $[a,b]$, et que sa somme

$$S(x) = \sum_{n=0}^{+\infty} u_n(x)$$

est continue par morceaux sur $[a,b]$. On a alors convergence de la série des intégrales $\int_a^b u_n(x)\,dx$ et

$$\int_a^b S(x)\,dx = \int_a^b \sum_{n=0}^{+\infty} u_n(x)\,dx = \sum_{n=0}^{+\infty} \int_a^b u_n(x)\,dx$$

Démonstration : on applique le théorème sur les suites aux sommes partielles de la série. Par linéarité de l'intégrale on a

$$\int_a^b S_n(x)\,dx = \sum_{k=0}^{n} \int_a^b u_k(x)\,dx$$

ce qui démontre aisément le résultat[7]. ■

[7]Nous verrons dans le chapitre d'intégration sur un intervalle quelconque des théorèmes plus puissants, où la notion de convergence uniforme n'intervient pas.

EXERCICE 11-2.22 Montrer que, pour x réel vérifiant $|x| < 1$

$$\ln(1+x) = \int_0^x \frac{dt}{1+t} = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$$

En utilisant l'exercice **11-2**.17 montrer que

$$\ln 2 = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}$$

(résultat déjà obtenu par une autre méthode dans le chapitre sur les séries).

Dérivation terme à terme.

PROPOSITION 11-2.23 Soient I un intervalle de $\mathbb{R}$ et $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de classe C^1 de I dans $\mathbb{F}$ tels que

1) Il existe $a \in I$ tel que $\sum u_n(a)$ converge

2) La série des dérivées $\sum u_n'$ converge uniformément sur tout segment inclus dans I

Alors la série $\sum u_n$ converge uniformément sur tout segment inclus dans I. Sa somme est une fonction de classe C^1 sur I et

$$\frac{d}{dx}\left(\sum_{n=0}^{+\infty} u_n(x)\right) = \sum_{n=0}^{+\infty} u_n'(x) \qquad \textbf{sur } I$$

Dans la pratique, on vérifie la convergence simple de la série de fonctions et la convergence uniforme (locale) de la série des dérivées pour être assuré du caractère C^1 de la somme de la série. On peut alors dériver la somme terme à terme.

On a évidemment un corollaire analogue à **11-1**.13 :

COROLLAIRE 11-2.24 Si $(u_n)_{n\in\mathbb{N}}$ est une suite de fonctions de classe C^p ($p \geqslant 1$) de I dans $\mathbb{F}$ telle que, pour $0 \leqslant k \leqslant p$, les séries $\sum u_n^{(k)}$ convergent simplement sur I, la convergence étant uniforme sur tout segment $\subset I$ pour la série $\sum u_n^{(p)}$ des dérivées d'ordre p. La somme de la série est alors de classe C^p sur I et, pour tout k entier vérifiant $1 \leqslant k \leqslant p$,

$$\frac{d^k}{dx^k}\left(\sum_{n=0}^{+\infty} u_n(x)\right) = \sum_{n=0}^{+\infty} u_n^{(k)}(x)$$

EXERCICE 11-2.25 Montrer que

$$S(x) = \sum_{n=1}^{+\infty} (-1)^n \frac{n}{n^2+x^2}$$

définit une fonction de classe C^∞ sur $\mathbb{R}$.

11-2.2 Exemples

EXEMPLE 11-2.26 On considère la série de fonctions de terme général

$$u_n(x) = n \cos^n x \sin x$$

Etudier la convergence simple sur $\left[0, \frac{\pi}{2}\right]$. Montrer qu'il y a convergence uniforme sur le segment $\left[\alpha, \frac{\pi}{2}\right]$ pour tout α vérifiant $0 < \alpha < \frac{\pi}{2}$. Calculer (en utilisant $u_n(x) = \cos x \, v_n'(x)$ avec $v_n(x) = -\cos^n x$) la somme

$$\sum_{n=0}^{+\infty} u_n(x)$$

Montrer qu'il n'y a pas convergence uniforme sur $\left[0, \frac{\pi}{2}\right]$.

EXEMPLE 11-2.27 Fonction Zeta de Riemann. On rappelle que, si $z \in \mathbb{C}$ et $a \in \mathbb{R}^{*+}$, on a par définition

$$a^z = e^{z.\ln a}$$

On pose

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$$

Dans un premier temps, nous supposons la variable s réelle : montrer que ζ est définie sur $]1, +\infty[$, et est de classe C^∞ sur cet intervalle, avec

$$\forall k \geqslant 1 \quad \forall s > 1 \quad \zeta^{(k)}(s) = \sum_{n=1}^{+\infty} \frac{(-\ln n)^k}{n^s}$$

Si on suppose la variable s complexe, montrer que le domaine de convergence de la série définissant ζ est exactement

$$\mathcal{D} = \{s \in \mathbb{C} \mid \operatorname{Re} s > 1\}$$

(sur ce domaine, montrer qu'il y a convergence absolue de la série; si $\operatorname{Re} s \leqslant 0$, il y a divergence grossière et si $0 < \operatorname{Re} s \leqslant 1$, étudier la série de terme général $\frac{1}{n^s} - \int_n^{n+1} \frac{dt}{t^s}$). Montrer que ζ est continue sur son domaine de définition.

EXEMPLE 11-2.28 Série de Riemann alternée. Montrer que

$$\zeta_1(s) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n^s}$$

définit une fonction C^∞ sur $]0, +\infty[$. Si la variable est complexe, montrer que ζ_1 est continue sur le demi-plan $\operatorname{Re} s > 0$ (dans $\operatorname{Re} s > 1$ ce n'est pas très difficile, sinon il faut adapter la technique des séries alternées au cas complexe, c'est à dire regrouper les termes deux par deux).

Essentiellement, dans les exemples précédents, c'est la convergence normale ou le théorème des séries alternées (éventuellement adapté au cas complexe) qui donnent des majorations uniformes des restes des séries, amenant à des résultats de convergence uniforme. Une autre technique est basée sur l'utilisation d'une transformation d'Abel pour prouver à la fois la convergence simple et obtenir des majorations uniformes des restes :

EXEMPLE 11-2.29 Si $\alpha > 0$ est un réel fixé, étudier la convergence simple et uniforme de la série de fonctions de la variable réelle x

$$\sum_{n=1}^{+\infty} \frac{e^{inx}}{n^\alpha}$$

Par périodicité, on peut se ramener à travaille sur $[-\pi, \pi]$. De plus, changer x en $-x$ conjugue le terme général. On se ramène donc à $x \in [0, \pi]$.

Le cas $\alpha > 1$ ne doit pas poser de problème. Si $0 < \alpha \leqslant 1$, il y a évidemment divergence de la série en $x = 0$. Montrons qu'il y a convergence simple sur $]0, \pi]$, avec convergence uniforme[8] sur $[\delta, \pi]$ pour tout $\delta > 0$. Pour cela, on prouve d'abord la convergence simple par utilisation du critère de Cauchy et une transformation d'Abel : si on pose

$$A_n(x) = \sum_{k=0}^{n} e^{ikx}$$

montrer que

$$\forall x \in]0, \pi] \quad |A_n(x)| \leqslant M(x) = \frac{2}{|1 - e^{ix}|}$$

Montrer que, pour $n \leqslant m$

$$\sum_{k=n}^{m} \frac{e^{ikx}}{k^\alpha} = \sum_{k=n}^{m} \frac{A_k(x) - A_{k-1}(x)}{k^\alpha}$$
$$= \frac{A_m(x)}{m^\alpha} - \frac{A_{n-1}(x)}{n^\alpha} + \sum_{k=n}^{m-1} A_k(x) \left[\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha}\right]$$

Ceci donne facilement

$$\left|\sum_{k=n}^{m} \frac{e^{ikx}}{k^\alpha}\right| \leqslant 2\frac{M(x)}{n^\alpha}$$

et permet de prouver, par utilisation du critère de Cauchy, la convergence simple de la série. En faisant tendre m vers $+\infty$, on obtient de plus une majoration du reste (d'ordre $n-1$)

$$\left|\sum_{k=n}^{+\infty} \frac{e^{ikx}}{k^\alpha}\right| \leqslant 2\frac{M(x)}{n^\alpha}$$

[8] Le théorème d'interversion des limites nous montre qu'il ne peut y avoir convergence uniforme sur $]0, \pi]$.

ce qui doit permettre de prouver une convergence uniforme sur $[\delta, \pi]$ pour $\delta > 0$. On en déduit en particulier que

$$S(x) = \sum_{n=1}^{+\infty} \frac{e^{inx}}{n}$$

est une fonction continue sur $]0, \pi]$ (donc sur $\mathbb{R} - 2\pi\mathbb{Z}$).

REMARQUE 11-2.30 Nous verrons dans la section sur les séries entières (les séries de Fourier fournissant un autre outil pour y arriver) que la somme S de cette série est de classe C^∞ sur $]0, \pi]$. Il est clair cependant que la série définissant S ne peut être dérivée terme à terme. Se souvenir que le théorème de dérivation ne donne qu'une condition **suffisante** assurant la régularité de la somme de la série.

11-3 Séries entières

11-3.1 Rayon de convergence

11-3.1.1 Définition

DÉFINITION 11-3.1 *On appelle série entière toute série de fonctions de terme général*

$$u_n : \mathbb{K} \to \mathbb{C} \quad u_n(z) = a_n z^n$$

avec $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *et où* $(a_n)_{n\in\mathbb{N}}$ *est une suite complexe arbitraire.*

La suite des sommes partielles de cette série est

$$S_n(z) = \sum_{k=0}^{n} a_k z^k$$

C'est donc une suite de polynômes, mais une suite très particulière : on passe de S_{n-1} à S_n en ajoutant un monôme de degré n (éventuellement le monôme nul).

L'usage veut que l'on note z la variable lorsqu'on a une série entière d'une variable complexe. Dans ce qui suit, nous utiliserons la lettre x pour représenter la variable lorsqu'elle est réelle. Une étude générale des séries entières sur $\mathbb{C}$ donnera évidemment, en se restreignant à $\mathbb{R}$, les résultats sur les séries entières réelles.

Notons également qu'une série de fonctions de la forme $\sum b_n z^{2n}$ ou plus généralement $\sum b_n z^{\varphi(n)}$ où φ est une application strictement croissante de $\mathbb{N}$ dans lui-même peut être considérée également comme une série entière : il suffit d'envisager la suite complexe $(a_n)_{n\in\mathbb{N}}$ avec $a_{\varphi(n)} = b_n$ et $a_p = 0$ si $p \notin \varphi(\mathbb{N})$. Il est alors évident que, si on note (S_n) la suite des sommes partielles de la série $\sum b_n z^{\varphi(n)}$ et (T_n) celle de la série $\sum a_n z^n$, on a

$$S_n = T_{\varphi(n)}$$

et que la première série converge si et seulement s'il en est de même pour la seconde (puisque, pour $\varphi(n) \leqslant p < \varphi(n+1)$, nous avons $T_p = T_{\varphi(n)}$).

Nous allons d'abord essayer de résoudre les problèmes suivants :

- Si $\sum a_n z^n$ est une série entière, que peut-on dire de son domaine de convergence

$$\Omega = \left\{ z \in \mathbb{K} \mid \sum a_n z^n \text{ converge} \right\}$$

- Quelles sont les propriétés de la somme de la série

$$S : \Omega \to \mathbb{C} \qquad z \mapsto S(z) = \sum_{k=0}^{+\infty} a_k z^k$$

11-3.1.2 Lemme d'Abel

On comprend bien que les séries géométriques vont jouer un rôle fondamental dans l'étude des séries entières : l'idée simple est souvent de majorer (ou minorer) le module du terme général par le terme général d'une série géométrique.

PROPOSITION 11-3.2 Soit $\sum a_k z^k$ une série entière. Si $\rho > 0$ est un réel tel que la suite $(a_n \rho^n)_{n \in \mathbb{N}}$ est bornée, alors la série entière converge absolument en tout $z \in \mathbb{K}$ vérifiant $|z| < \rho$ (donc dans $]-\rho, \rho[$ dans le cas réel, dans le disque ouvert $D(0, \rho[$ de centre 0 et de rayon ρ dans le cas complexe).

Démonstration : il suffit de remarquer que, pour $|z| < \rho$

$$|a_n z^n| = |a_n \rho^n| . \left| \frac{z}{\rho} \right|^n \leqslant M. \left(\frac{|z|}{|\rho|} \right)^n$$

(où M est un majorant de la suite $(a_n \rho^n)_{n \in \mathbb{N}}$) ce qui majore le module du terme général de la série par le terme général d'une série géométrique convergente, puisque de raison < 1. ■

COROLLAIRE 11-3.3 S'il existe $z_0 \neq 0$ tel que la série $\sum a_k z_0^k$ converge, la série entière $\sum a_k z^k$ est absolument convergente en tout $z \in \mathbb{K}$ vérifiant $|z| < |z_0|$. De même, si la série $\sum a_k z_0^k$ est divergente, la série entière $\sum a_k z^k$ est divergente (grossièrement) en tout $z \in \mathbb{K}$ vérifiant $|z| > |z_0|$.

Démonstration : conséquence immédiate du lemme d'Abel. Dans le cas de la divergence de $\sum a_k z_0^k$, on est assuré que, pour $|z| > |z_0|$, la suite $(a_n z^n)_{n \in \mathbb{N}}$ ne peut être bornée, et *a fortiori* tendre vers 0. ■

11-3.1.3 Rayon de convergence

DÉFINITION 11-3.4 (ET THEOREME) *Si $\sum a_k z^k$ est une série entière, on appelle rayon de convergence de cette série le nombre $R \in \mathbb{R}^+ \cup \{+\infty\}$ défini par*

$$R = \sup \{\rho \in \mathbb{R}^+ \mid \textit{la suite } (a_n \rho^n) \textit{ est bornée}\}$$

La série $\sum a_k z^k$ est absolument convergente pour z vérifiant $z = 0$ ou $|z| < R$ et grossièrement divergente si $|z| > R$.

Démonstration : Remarquons que

$$I = \{\rho \in \mathbb{R}^+ \mid \text{la suite } (a_n \rho^n) \text{ est bornée}\}$$

n'est pas vide puisqu'il contient 0. De plus, une majoration analogue à celle utilisée dans la démonstration du lemme d'Abel montre que

$$\rho \in I \Rightarrow \forall \rho' \in [0, \rho] \quad \rho' \in I$$

Il en résulte que I est un intervalle de $\mathbb{R}^+$. S'il n'est pas borné, c'est $[0, +\infty[$ et dans ce cas $R = +\infty$. S'il est borné, R étant sa borne supérieure, on peut avoir $I = [0, R]$ ou $I = [0, R[$.

- Si $R = +\infty$, pour tout $z \in \mathbb{K}$ on a $\rho = |z| + 1 \in I$, et d'après le lemme d'Abel, la série $\sum a_k z^k$ converge absolument.

- Si $R = 0$, pour $z \neq 0$ la suite $(a_n z^n)$ n'est pas bornée, et la série $\sum a_k z^k$ diverge grossièrement.

- Si $0 < R < +\infty$, si $z \in \mathbb{K}$ vérifie $|z| < R$, alors $\rho = \dfrac{|z| + R}{2} \in I$ et vérifie $|z| < \rho$. Le lemme d'Abel entraîne la convergence absolue de $\sum a_k z^k$. Si $|z| > R$, alors $|z| \notin I$ et la suite $(a_n z^n)_{n \in \mathbb{N}}$ n'est pas bornée, ce qui entraîne la divergence grossière de $\sum a_k z^k$. ■

En résumé :

- Lorsque $R = +\infty$, la série entière est absolument convergente en tout point de $\mathbb{K}$. C'est ce que nous avons vu par exemple pour la série définissant l'exponentielle

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

- Lorsque $R = 0$, la série est partout divergente, sauf en $z = 0$. C'est le cas pour la série

$$\sum_{n=0}^{+\infty} n!\, z^n$$

- Lorsque $0 < R < +\infty$, le disque ouvert $D(0, R[$ est appelé **disque ouvert de convergence de la série entière** (on parlera évidemment d'intervalle ouvert de convergence $]-R, R[$ dans le cas réel). C'est un domaine de convergence absolue de la série. **Sur le complémentaire du disque fermé** $D(0, R]$ (de $[-R, R]$ dans le cas réel) **il y a divergence grossière de la série.**

DÉFINITION 11-3.5 *Si R est le rayon de convergence de la série $\sum a_k z^k$, le disque ouvert $D(0, R[$ (égal à $\mathbb{C}$ si $R = +\infty$, à $\emptyset$ si $R = 0$) est appelé disque ouvert de convergence de la série entière. Lorsque $0 < R < +\infty$, le cercle $C(0, R)$ de centre 0 et de rayon R est appelé cercle d'incertitude de la série entière.*

On dit cercle d'incertitude, car l'étude précédente ne donne pas d'information sur la nature de la série pour $|z| = R$. Par exemple les séries entières

$$\sum_{n=1}^{+\infty} \frac{z^n}{n^2}, \sum_{n=1}^{+\infty} \frac{z^n}{n} \text{ et } \sum_{n=1}^{+\infty} z^n$$

ont toutes un rayon de convergence égal à 1 (dans les trois cas, l'intervalle I considéré dans la démonstration précédente vaut $[0, 1]$). Pour la première série, il y a convergence (absolue) en tout point du cercle unité. La seconde est semi-convergente (*cf.* exemple **11-2**.29) pour tout z de module 1 distinct de 1 et est divergente en $z = 1$. La troisième est évidemment divergente en tout point du cercle d'incertitude.

11-3.1.4 Exemples de détermination

Déterminer le rayon de convergence d'une série entière, c'est essentiellement "peser" le terme général $a_n z^n$, voir quand il forme une suite bornée, ce qui revient essentiellement à comparer la suite a_n à une suite géométrique.

Un premier résultat est évident à partir de la définition de R comme borne supérieure de

$$\{\rho \in \mathbb{R}^+ \mid \text{la suite } (a_n \rho^n) \text{ est bornée}\}$$

PROPOSITION 11-3.6 Si $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites complexes avec

$$a_n \underset{n\to+\infty}{\sim} b_n$$

les séries entières $\sum a_k z^k$ et $\sum b_k z^k$ ont même rayon de convergence[9].

[9]L'équivalence $a_n z^n \underset{n\to+\infty}{\sim} b_n z^n$ montre que les domaines de convergence absolue des deux séries sont les mêmes : ce sont $D(0, R[$ ou $D(0, R]$ pour les deux séries. Par contre les domaines de convergence peuvent être distincts. Par exemple

$$\sum_{n=1}^{+\infty} \left(\frac{(-1)^n}{\sqrt{n}} + \frac{1}{n}\right) z^n \text{ et } \sum_{n=1}^{+\infty} \frac{(-1)^n}{\sqrt{n}} z^n$$

ont des natures différentes en $z = 1$.

D'autres caractérisations du rayon de convergence sont aussi utilisables :

EXERCICE 11-3.7 Montrer que le rayon de convergence R d'une série $\sum a_k z^k$ vérifie

$$R = \sup\left\{|z| \mid z \in \mathbb{C} \text{ et } \lim_{n \to +\infty} a_n z^n = 0\right\}$$
$$R = \sup\left\{|z| \mid z \in \mathbb{C} \text{ et } \sum a_n z^n \text{ converge}\right\}$$
$$R = \sup\left\{|z| \mid z \in \mathbb{C} \text{ et } \sum a_n z^n \text{ converge absolument}\right\}$$

(les ensembles de complexes considérés sont en général distincts, mais les bornes supérieures des modules de leurs éléments sont égales).

Les propriétés suivantes découlent aussi immédiatement des études précédentes :

- Si R_a et R_b sont les rayons de convergence respectifs des séries entières $\sum a_k z^k$ et $\sum b_k z^k$

$$a_n \underset{n \to +\infty}{=} O(b_n) \Rightarrow R_a \geqslant R_b$$

- Si $\sum a_k z_0^k$ diverge alors $R_a \leqslant |z_0|$.
- Si $\sum a_k z_0^k$ converge alors $R_a \geqslant |z_0|$.

EXEMPLE 11-3.8 Déterminer les rayons de convergence des séries entières

$$\sum_{n=1}^{+\infty} \frac{z^n}{n^\alpha},\ \sum_{n=1}^{+\infty} \left(\sin \frac{1}{\sqrt{n}}\right) z^n,\ \sum_{n=1}^{+\infty} \frac{\sin n}{n} z^n,\ \sum_{n=1}^{+\infty} \frac{(n+1)^{n-1}}{n!} z^n$$

Les deux premières séries ne devraient pas poser de problème. La majoration

$$\left|\frac{\sin n}{n}\right| \leqslant 1$$

montre que le rayon de convergence de la troisième série est au moins égal à celui de la série $\sum z^n$ soit 1. Si l'on montre que la suite $(\sin n)_{n\in\mathbb{N}}$ ne tend pas vers 0, on obtient facilement que

$$|z| > 1 \Rightarrow \left(\frac{\sin n}{n} z^n\right)_{n\in\mathbb{N}} \quad \text{non bornée}$$

ce qui permet de conclure. Pour la quatrième série, chercher un équivalent du coefficient (Stirling) ou, plus simplement ici, utiliser la règle de D'Alembert (voir proposition suivante).

EXERCICE 11-3.9 Pour les première, seconde et quatrième séries qui précédent, faire une étude de convergence sur le cercle d'incertitude.

Une utilisation de la règle de D'Alembert permet **parfois** de déterminer le rayon de convergence d'une série entière :

PROPOSITION 11-3.10 Si $\sum a_k\, z^k$ est une série entière dont le coefficient a_n ne s'annule pas à partir d'un certain rang, et si la limite $l = \lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right|$ existe dans $\mathbb{R}^+ \cup \{+\infty\}$, le rayon de convergence de cette série vérifie alors

$$R = \frac{1}{l}$$

(avec les conventions $\frac{1}{0} = +\infty$ et $\frac{1}{+\infty} = 0$).

Démonstration : On s'intéresse au domaine de convergence absolue de la série entière (rappelons que la règle de D'Alembert est une méthode d'étude d'absolue convergence). Si $u_n(z) = a_n\, z^n$ est le terme général de la série, on a, pour $z \neq 0$,

$$\left|\frac{u_{n+1}(z)}{u_n(z)}\right| = \left|\frac{a_{n+1}}{a_n}\right| . |z| \underset{n\to+\infty}{\to} l.|z|$$

On sait que lorsque cette limite est < 1, la série converge absolument. Lorsqu'elle est > 1, la série diverge grossièrement. On en déduit aisément

$$R = \frac{1}{l} \quad \blacksquare$$

Attention ! Cette règle semble d'un usage si commode qu'elle est souvent mal utilisée : le rayon de convergence de $\sum a_k\, z^k$ existe toujours, ce qui n'est pas nécessairement le cas de $\lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right|$. Il faut bien se garder d'appliquer cette règle "à l'envers" et d'affirmer : "comme le rayon de convergence vaut R, la limite de $\left|\frac{a_{n+1}}{a_n}\right|$ est $\frac{1}{R}$". On peut rendre cet énoncé correct en disant "si la limite de $\left|\frac{a_{n+1}}{a_n}\right|$ existe, c'est forcément $\frac{1}{R}$", mais on ne peut dire plus.

Il serait stupide de vouloir appliquer la règle de D'Alembert à la troisième série de l'exemple **11-3**.8. Ce serait bien plus astucieux pour la quatrième série, où la présence de puissance et de factorielles permet d'espérer une simplification intéressante lors du calcul de $\left|\frac{a_{n+1}}{a_n}\right|$.

De plus, il semble préférable, au cas par cas, de refaire la démonstration qui précède lorsque l'on veut appliquer cette "recette". Ainsi, pour une série entière de la forme

$$\sum a_n\, z^{2n+5}$$

pour laquelle $\left|\frac{a_{n+1}}{a_n}\right| \to l$, on aura cette fois, avec des notations évidentes,

$$\left|\frac{u_{n+1}(z)}{u_n(z)}\right| = \left|\frac{a_{n+1}}{a_n}\right| |z|^2 \underset{n\to+\infty}{\to} l\; |z|^2$$

ce qui donnera

$$R = \frac{1}{\sqrt{l}} \text{ et non pas } R = \frac{1}{l}$$

EXERCICE 11-3.11 (le résultat de cet exercice est fondamental) : montrer que le rayon de convergence d'une série entière $\sum a_k z^k$ est non nul si et seulement s'il existe un réel $b > 0$ tel que

$$a_n \underset{n \to +\infty}{=} O(b^n)$$

donc si la suite des coefficients est dominée par une suite géométrique.

EXERCICE 11-3.12 (Formule d'Hadamard) Si on note $\overline{\lim}\sqrt[n]{|a_n|}$ la plus grande valeur d'adhérence (dans $\mathbb{R}^+ \cup \{+\infty\}$) de la suite $\left(\sqrt[n]{|a_n|}\right)_{n\in\mathbb{N}}$, montrer que

$$R = \frac{1}{\overline{\lim}\sqrt[n]{|a_n|}}$$

11-3.1.5 Rayon de convergence et opérations algébriques

Sommes et combinaisons linéaires

DÉFINITION 11-3.13 *Si $\sum a_n z^n$ et $\sum b_n z^n$ sont deux séries entières, on appelle somme de ces deux séries la série entière*

$$\sum c_n z^n \quad \text{avec } \forall n \in \mathbb{N} \quad c_n = a_n + b_n$$

De même, une série entière de la forme

$$\sum (\alpha\, a_n + \beta\, b_n)\, z^n$$

où α et β sont deux complexes donnés sera dite combinaison linéaire des séries entières $\sum a_n z^n$ et $\sum b_n z^n$.

PROPOSITION 11-3.14 On ne change pas le rayon de convergence d'une série entière en la multipliant par un scalaire non nul. Si $\sum a_k z^k$ et $\sum b_k z^k$ sont deux séries entières de rayons de convergence respectifs R_a et R_b, le rayon de convergence R_c de leur somme vérifie

$$R_c \geqslant \inf(R_a, R_b)$$

Il y a égalité lorsque $R_a \neq R_b$. Lorsque $R_a = R_b$, il peut y avoir inégalité stricte. Si S_a, S_b et S_c représentent les sommes de ces séries, on a évidemment

$$|z| < \inf(R_a, R_b) \Rightarrow S_c(z) = S_a(z) + S_b(z)$$

et plus généralement

$$\forall \alpha \in \mathbb{C} \quad |z| < \inf(R_a, R_b) \Rightarrow \sum_{n=0}^{+\infty} (\alpha\, a_n + b_n)\, z^n = \alpha \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$$

Démonstration : Si $z \in \mathbb{C}$ vérifie $|z| < \inf(R_a, R_b)$, les séries $\sum a_k z^k$ et $\sum b_k z^k$ sont toutes deux absolument convergentes. Il en est donc de même de la série somme, puisque

$$|c_n z^n| \leqslant |a_n z^n| + |b_n z^n|$$

Ceci prouve bien que $R_c \geqslant \inf(R_a, R_b)$. Lorsque les deux rayons sont distincts, avec par exemple $R_a < R_b$, pour ρ vérifiant $R_a < \rho < R_b$, la suite $((a_n + b_n)\rho^n)_{n\in\mathbb{N}}$ n'est pas bornée (somme d'une suite bornée et d'une suite non bornée). Il en résulte que R_c ne peut être supérieur à R_a et on a donc $R_c = R_a$. Lorsque $R_a = R_b$, il peut y avoir des "compensations" : par exemple, les séries

$$\sum_{n=1}^{+\infty} \left(\frac{1}{n} + \frac{1}{n!}\right) z^n \text{ et } \sum_{n=1}^{+\infty} \left(-\frac{1}{n} + \frac{1}{n!}\right) z^n$$

ont toutes deux un rayon de convergence égal à 1. Le rayon de leur somme est cependant infini. ■

Un cas particulier est important, car il se rencontre très souvent en pratique : il s'agit de deux séries entières $\sum a_k z^k$ et $\sum b_k z^k$ avec

$$\forall n \in \mathbb{N} \quad a_n b_n = 0$$

(on dit parfois que ces deux séries sont "disjointes" : par exemple, la première série ne fait intervenir que des monômes pairs, la seconde des monômes impairs). On a alors dans ce cas toujours

$$R_c = \inf(R_a, R_b)$$

puisque

$$\forall n \quad |c_n z^n| = |a_n z^n| + |b_n z^n|$$

et donc la suite $(c_n z^n)$ ne peut être bornée que si les deux suites $(a_n z^n)$ et $(b_n z^n)$ le sont.

EXERCICE 11-3.15 Si $\sum a_k z^k$ et $\sum b_k z^k$ ont pour rayons de convergence respectifs R_a et R_b, trouver le rayon de convergence de la série $\sum d_k z^k$, avec

$$\forall n \in \mathbb{N} \quad d_{2n} = a_n \text{ et } d_{2n+1} = b_n$$

Produit de Cauchy.

PROPOSITION 11-3.16 Si $\sum a_n z^n$ et $\sum b_n z^n$ sont deux séries entières, le produit de Cauchy de ces deux séries est une série entière (dite série produit des deux séries entières)

$$\sum c_n z^n \quad \textbf{avec } \forall n \in \mathbb{N} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{\substack{(p,q)\in\mathbb{N}^2 \\ p+q=n}} a_p b_q$$

Démonstration : C'est simplement la définition du produit de Cauchy :

$$c_n\, z^n = \sum_{k=0}^{n} \left(a_k\, z^k\right) . \left(b_{n-k}\, z^{n-k}\right) \quad \blacksquare$$

Nous avons vu lors de l'étude des familles sommables que le produit de Cauchy de deux séries absolument convergentes était une série absolument convergente, dont la somme est le produit des sommes des deux séries. Comme une série entière est absolument convergente sur son disque ouvert de convergence, on obtient immédiatement :

THÉORÈME 11-3.17 Si $\sum a_n\, z^n$ et $\sum b_n\, z^n$ sont deux séries entières de rayons de convergence respectifs R_a et R_b, le rayon de convergence R_c de leur produit vérifie

$$R_c \geqslant \inf(R_a, R_b)$$

et, pour tout z complexe vérifiant $|z| < \inf(R_a, R_b)$, on a

$$\sum_{n=0}^{+\infty} c_n\, z^n = \left(\sum_{n=0}^{+\infty} a_n\, z^n\right) \left(\sum_{n=0}^{+\infty} b_n\, z^n\right)$$

On remarquera que, contrairement à ce qui se passe pour la somme, il peut y avoir inégalité stricte même lorsque $R_a \neq R_b$: si on prend par exemple

$$\forall n \quad a_n = 1 + \frac{1}{3^{n+1}} \quad \text{et} \quad \forall n \geqslant 1 \quad b_n = -\frac{1}{2^{n+1}} \text{ avec } b_0 = \frac{1}{2}$$

on vérifie aisément que $R_a = 1$, $R_b = 2$ et $R_c = 3$.

EXERCICE 11-3.18 Calculer la somme de chacune de ces trois séries sur leurs disques ouverts de convergence, pour comprendre où se situe la "compensation".

Dérivation formelle.

DÉFINITION 11-3.19 *Si $\sum a_n\, z^n$ est une série entière, on appelle dérivée formelle de cette série la série entière*

$$\sum_{n\geqslant 1} n\, a_n\, z^{n-1}$$

En itérant le procédé, on obtient la série dérivée formelle d'ordre $p \geqslant 2$

$$\sum_{n\geqslant p} n\,(n-1)\cdots(n-p+1)\, a_n\, z^{n-p} = \sum_{n\geqslant 0} \frac{(n+p)!}{n!} a_{n+p}\, z^n$$

la seconde écriture étant obtenue par décalage des indices.

Lors de la dérivation formelle, le fait de multiplier le terme général par n est de peu d'importance devant les suites géométriques $n \mapsto z^n$:

THÉORÈME 11-3.20 Une série entière et sa dérivée formelle ont même rayon de convergence.

Démonstration : soit $\sum a_n z^n$ une série entière de rayon de convergence R. Notons R' le rayon de convergence de sa dérivée formelle. Si $\rho \geqslant 0$ est un réel tel que la suite $(na_n\rho^{n-1})_{n\in\mathbb{N}}$ est borné, il en est évidemment de même de la suite $(a_n\rho^n)_{n\in\mathbb{N}}$, ce qui entraîne

$$R \geqslant R'$$

Si $R = 0$, on a immédiatement $R' = 0$. Sinon, prenons un réel ρ vérifiant $0 < \rho < R$, et posons $\rho_1 = \dfrac{R+\rho}{2}$ si $R < +\infty$, $\rho_1 = 2\rho$ si $R = +\infty$. Par définition de R, la suite $(a_n\rho_1^n)_{n\in\mathbb{N}}$ est bornée. Comme

$$na_n\rho^{n-1} = \frac{1}{\rho}(a_n\rho_1^n)\cdot\left(n\left(\frac{\rho}{\rho_1}\right)^n\right)$$

est produit de deux suites bornées, on a $\rho \leqslant R'$. Comme ceci est valable pour tout $\rho < R$, on a aussi

$$R \leqslant R' \quad \blacksquare$$

COROLLAIRE 11-3.21 La série $\sum a_n z^n$ a même rayon de convergence que toutes ses dérivées formelles d'ordre quelconque.

COROLLAIRE 11-3.22 La série entière obtenue par intégration terme à terme

$$\sum \frac{a_n}{n+1} z^{n+1}$$

a même rayon de convergence que la série $\sum a_n z^n$.

Démonstration : C'est évident, puisque la dérivée formelle de cette nouvelle série est la série de départ. $\blacksquare$

EXERCICE 11-3.23 Montrer que, plus généralement, si Q est une fraction rationnelle n'ayant pas de pôle entier

$$\sum a_n z^n \quad \text{et} \quad \sum Q(n)\, a_n z^n$$

ont même rayon de convergence.

REMARQUE 11-3.24 Si les rayons de convergence ne changent pas par intégration ou dérivation terme à terme, les domaines de convergence peuvent être différents : en un point du cercle d'incertitude, une série entière et sa série dérivée peuvent ne pas avoir le même comportement. Voir par exemple les séries

$$\sum_{n=1}^{+\infty} \frac{z^n}{n^2}, \quad \sum_{n=1}^{+\infty} \frac{z^{n-1}}{n} \text{ et } \sum_{n=2}^{+\infty} \frac{n-1}{n} z^{n-2}$$

Pour le moment, nous n'avons pas relié la somme de la série dérivée à la somme de la série de départ. On imagine bien qu'une utilisation des théorèmes vus à la section **11-2.1.4** permettra d'en dire plus. C'est ce que nous allons voir à présent.

11-3.2 Propriétés de la somme d'une série entière

Nous étudions ici les propriétés de régularité (continuité, dérivabilité, etc...) de la somme d'une série entière sur son domaine de convergence. L'outil de base est évidemment la convergence uniforme. Nous énonçons les théorèmes dans le cas d'une variable complexe. En remplaçant "disque ouvert de convergence" par "intervalle ouvert de convergence", on obtient les mêmes résultats lorsque la variable est réelle.

11-3.2.1 Domaines de convergence uniforme

THÉORÈME 11-3.25 Si $\sum a_k z^k$ est une série entière de rayon de convergence $R > 0$, cette série converge normalement et donc uniformément sur tout disque fermé centré en 0 inclus dans le disque ouvert de convergence.

Démonstration : Soit $D(0,\rho] \subset D(0,R[$ un disque fermé. Si ρ_1 est choisi vérifiant $0 < \rho < \rho_1$, on a

$$\forall z \in D(0,\rho] \quad \forall n \quad |a_n z^n| \leqslant |a_n \rho_1^n| \cdot \left(\frac{\rho}{\rho_1}\right)^n \leqslant M\left(\frac{\rho}{\rho_1}\right)^n$$

si M est un majorant de la suite (bornée) $(a_n\rho_1^n)_{n\in\mathbb{N}}$. Comme $\dfrac{\rho}{\rho_1} < 1$, la série géométrique majorante est convergente. ∎

ATTENTION : IL N'Y A PAS, EN GENERAL, CONVERGENCE UNIFORME SUR LE DISQUE OUVERT DE CONVERGENCE.

Le contre-exemple le plus simple est donné par la série géométrique $\sum z^n$, pour laquelle $R = 1$, et qui ne converge évidemment pas uniformément sur $D(0,1[$, puisque son terme général ne tend pas uniformément vers 0.

COROLLAIRE 11-3.26 Une série entière de rayon de convergence $R > 0$ converge uniformément sur tout compact inclus dans le disque ouvert de convergence.

Démonstration : Si $K \subset D(0,R[$ est un compact, la fonction continue $z \mapsto |z|$ atteint son maximum sur K en au moins un point z_0 qui vérifie $|z_0| < R$. On a alors

$$\forall z \in K \quad |z| \leqslant |z_0| < R$$

et donc $K \subset D(0,|z_0|] \subset D(0,R[$. La convergence normale sur $D(0,|z_0|]$ entraîne la convergence normale sur K. ∎

11-3.2.2 Continuité

THÉORÈME 11-3.27 Si $\sum a_k z^k$ est une série entière de rayon de convergence $R > 0$, sa somme

$$S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est une fonction continue sur $D(0,R[$.

Démonstration : Si $z_0 \in D(0, R[$, la continuité de S en z_0 se prouve en démontrant que la convergence de la série est uniforme sur un voisinage de z_0, puisque le terme général de la série est une fonction continue en z_0. Si on choisit ρ avec $|z_0| < \rho < R$, $D(0, \rho]$ est un tel voisinage d'après le théorème **11-3**.25. ■

COROLLAIRE 11-3.28 (Unicité d'un développement en série entière) Soit $\sum a_k z^k$ une série entière de rayon de convergence $R > 0$ et

$$S : D(0, R[\to \mathbb{C} \qquad z \mapsto S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

Si $\sum b_k z^k$ est une autre série entière qui représente S au voisinage de 0, c'est-à-dire telle que

$$\exists \alpha > 0 \quad \forall z \in D(0, \alpha[\quad S(z) = \sum_{n=0}^{+\infty} b_n z^n$$

alors

$$\forall n \in \mathbb{N} \qquad a_n = b_n$$

Démonstration : on a évidemment $a_0 = b_0 = S(0)$. Si on suppose que les suites (a_n) et (b_n) sont distinctes, posons

$$p = \inf \{n \in \mathbb{N}^* \mid a_n \neq b_n\}$$

On a alors

$$\forall z \in D(0, \alpha[\quad \sum_{n=p}^{+\infty} b_n z^n = \sum_{n=p}^{+\infty} a_n z^n$$

ce qui peut s'écrire aussi, en divisant par z^p

$$0 < |z| < \alpha \Rightarrow \sum_{n=p}^{+\infty} (a_n - b_n) z^{n-p} = 0$$

La série entière $\sum_{n \geqslant p} (a_n - b_n) z^{n-p}$, (dont le rayon de convergence est par hypothèse au moins égal à α) est continue à l'origine. L'égalité précédente est donc aussi valable en $z = 0$, ce qui donne $a_p = b_p$ et est contradictoire[10] avec les hypothèses faites sur les suites (a_n) et (b_n). ■

[10] Si on reprend attentivement cette démonstration, il suffirait de supposer que l'égalité

$$S(z) = \sum_{n=0}^{+\infty} b_n z^n$$

ait lieu en tout point d'une suite de complexes non stationnaire tendant vers 0 pour arriver à la même conclusion. C'est une autre manière de considérer le principe des zéros isolés pour une fonction analytique, résultat sur lequel nous reviendrons ultérieurement.

COROLLAIRE 11-3.29 Si $\sum a_k z^k$ est une série entière de rayon de convergence $R > 0$ et si sa somme

$$S : D(0, R[\to \mathbb{C} \qquad z \mapsto S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est une fonction paire alors

$$\forall n \in \mathbb{N} \qquad a_{2n+1} = 0$$

Démonstration : Il suffit de comparer les séries entières représentant $S(z)$ et $S(-z)$. De même, si on supposait S impaire, ce serait la suite $(a_{2n})_{n\in\mathbb{N}}$ qui serait identiquement nulle. ■

COROLLAIRE 11-3.30 Si $\sum a_k z^k$ est une série entière de rayon de convergence $R > 0$ et

$$S : D(0, R[\to \mathbb{C} \qquad z \mapsto S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est sa somme, S admet à l'origine un développement limité à tout ordre, obtenu en tronquant la série

$$S(z) \underset{z\to 0}{=} \sum_{k=0}^{p} a_k z^k + O\left(z^{p+1}\right)$$

Démonstration : On a en effet, pour $|z| < R$

$$S(z) = \sum_{k=0}^{p} a_k z^k + z^{p+1} \left(\sum_{n=0}^{+\infty} a_{n+p+1} z^n \right)$$

La somme de la série $\sum_{n=0}^{+\infty} a_{n+p} z^n$ étant continue en 0 est par conséquent bornée au voisinage de 0. ■

REMARQUE 11-3.31 ATTENTION : Les écritures

$$S(z) = \sum_{n=0}^{+\infty} a_n z^n \quad , |z| < R$$

$$\forall p \quad S(z) \underset{z\to 0}{=} \sum_{k=0}^{p} a_k z^k + O\left(z^{p+1}\right)$$

n'ont pas du tout la même utilité. La première donne une représentation de la fonction S dans tout le disque de centre 0 et de rayon R. Elle permet notamment d'obtenir des valeurs approchées de $S(z_0)$ dès que $|z_0| < R$. La seconde ne donne que le comportement asymptotique de $S(z)$ pour $z \to 0$, par exemple pour lever des indéterminations, trouver des équivalents, obtenir des majorations sur un voisinage de 0 sans avoir d'information sur la taille de ce voisinage[11].

[11]Nous reviendrons ultérieurement sur le fait qu'il est beaucoup plus fort de supposer qu'une fonction S puisse s'écrire sur un voisinage de 0 comme somme d'une série entière que de supposer qu'elle possède un développement limité à tout ordre !

REMARQUE 11-3.32 Aussi curieux que cela puisse paraître, il n'y a pas, dans le cas complexe, de résultat général de continuité de la somme d'une série entière sur son domaine de convergence. La somme de la série est continue sur le disque ouvert de convergence. Elle peut ne pas l'être en un point de convergence situé sur le cercle d'incertitude. Il y a cependant une propriété de continuité "radiale" que nous développons dans la section suivante.

11-3.2.3 Exercice : étude en un point du cercle d'incertitude

Le problème est le suivant : une série entière $\sum a_k z^k$ a un rayon de convergence R vérifiant $0 < R < +\infty$, et on suppose que la série $\sum a_k z_0^k$ converge, pour un point z_0 vérifiant $|z_0| = R$. On étudie les propriétés de continuité de S

$$S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

en z_0. Il est deux cas, que l'on rencontre assez souvent en pratique, où l'étude est simple :

- **Cas de la convergence absolue :**

  Si la série $\sum a_n z_0^n$ converge absolument, c'est-à-dire si

  $$\sum_{n=0}^{+\infty} |a_n| R^n < +\infty$$

  la convergence de la série est alors normale sur $D(0,R]$ (qui est alors le domaine de convergence de la série), et S est continue sur $D(0,R]$, donc en particulier en z_0. Il n'y a par exemple aucune difficulté à prouver la continuité en $z = 1$ (ou $z = e^{i\theta}$, $\theta \in \mathbb{R}$) de

  $$S : D(0,1] \to \mathbb{C} \qquad z \mapsto \sum_{n=1}^{+\infty} \frac{z^n}{n^2}$$

- **Cas d'une variable réelle et de l'utilisation du théorème des séries alternées :**

  On suppose à présent $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ et

  $$S(x) = \sum_{n=0}^{+\infty} a_n x^n \quad \text{définie sur }]-R,R]$$

  la convergence en tout point de $]0,R]$ étant assurée par **les hypothèses du théorème des séries alternées** (signe alterné pour le terme général, dont la valeur absolue est décroissante à partir d'un rang n_0 **indépendant de** x et tend vers 0). La convergence de la série est alors uniforme sur $[0,R]$ (ce qui assure bien la continuité de S en R) puisque

  $$\forall x \in [0,R] \quad \forall n \geqslant n_0 \quad |R_n(x)| = \left|\sum_{k=n+1}^{+\infty} a_k x^k\right| \leqslant \left|a_{n+1} x^{n+1}\right| \leqslant |a_{n+1}| R^{n+1}$$

majoration prouvant la convergence uniforme des restes de la série vers 0. Il y a bien convergence uniforme sur $[0, R]$.

C'est ainsi qu'ayant prouvé (*cf.* exercice **11-2**.22) que

$$\forall x \in]-1,1[\quad \ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$$

on peut en déduire

$$\ln 2 = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}$$

L'étude précédente est un cas particulier d'un résultat de continuité "radiale", dû à Abel. La démonstration utilise une transformation ... d'Abel comme outil principal !

EXERCICE 11-3.33 (ET PROPOSITION) Si $\sum a_k z^k$ a un rayon de convergence R tel que $0 < R < +\infty$ et si la série est convergente en z_0 vérifiant $|z_0| = R$, la somme de la série est continue en restriction au rayon $[0, z_0]$.

En posant, pour $t \in [0,1]$

$$F(t) = \sum_{n=0}^{+\infty} a_n (t z_0)^n = \sum_{n=0}^{+\infty} u_n t^n$$

où $u_n = a_n z_0^n$ est, par hypothèse, le terme général d'une série convergente, on se ramène en fait au cas particulier où $[0, z_0] = [0, 1]$. Pour prouver la continuité de F en 1, on montre la convergence uniforme de la série sur $]0, 1[$ en prouvant la convergence uniforme des restes vers 0. Comme la série de terme général u_n est convergente, la suite des restes

$$r_n = \sum_{k=n}^{+\infty} u_k$$

est définie et tend vers 0. On a de plus

$$u_n = r_n - r_{n+1}$$

ce qui va permettre de trouver une autre expression de

$$R_n(t) = \sum_{k=n}^{+\infty} u_k t^k = \sum_{k=n}^{+\infty} (r_k - r_{k+1}) t^k \quad \text{avec } t \in]0, 1[$$

Pour $n \leqslant p$, on a en effet

$$\sum_{k=n}^{p} (r_k - r_{k+1}) t^k = r_n t^n - r_{p+1} t^p + \sum_{k=n+1}^{p} r_k \left(t^k - t^{k-1}\right)$$

Faisant tendre p vers $+\infty$, comme $0 < t < 1$ et puisque la suite (r_n) tend vers 0, on obtient :

$$\forall t \in]0,1[\quad R_n(t) = r_n t^n + \sum_{k=n+1}^{+\infty} r_k \left(t^k - t^{k-1}\right)$$

(la majoration $\left|r_k\left(t^k - t^{k-1}\right)\right| \leqslant \sup_{\mathbb{N}} |r_n| \times \left(t^{k-1} - t^k\right)$ prouve la convergence absolue de la série du second membre).

Si $\varepsilon > 0$ est fixé arbitrairement, on peut trouver un rang n_0 tel que

$$k \geqslant n_0 \Rightarrow |r_k| \leqslant \varepsilon$$

On a alors, par inégalité triangulaire

$$\forall n \geqslant n_0 \quad \forall t \in]0,1[\quad |R_n(t)| \leqslant \varepsilon t^n + \sum_{k=n+1}^{+\infty} \varepsilon \left(t^{k-1} - t^k\right) = 2\varepsilon t^n \leqslant 2\varepsilon$$

ce qui prouve bien la convergence uniforme des restes vers 0. ■

11-3.2.4 Intégration

On s'intéresse ici évidemment à des séries entières d'une variable réelle (mais la suite (a_n) des coefficients peut être complexe) :

THÉORÈME 11-3.34 Si $\sum a_k x^k$ est une série entière de rayon de convergence $R > 0$, la fonction

$$f :]-R, R[\to \mathbb{C} \quad f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

est continue sur $]-R, R[$ et

$$\forall x \in]-R, R[\quad \int_0^x f(t)\, dt = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

Démonstration : On sait déjà que le rayon de convergence de la série obtenue par intégration formelle est égal à R, ce qui assure l'existence de la somme de la série du second membre. Cette existence et la valeur de cette somme peuvent aussi s'obtenir par le théorème d'intégration terme à terme : la série de fonctions

$$\sum a_n t^n$$

est une série de fonctions continues qui converge uniformément car normalement sur $[0, x]$ (ou $[x, 0]$). La somme est donc continue sur ce segment, et l'intégration terme à terme est possible, ce qui donne

$$\int_0^x \left(\sum_{n=0}^{+\infty} a_n t^n\right) dt = \sum_{n=0}^{+\infty} \int_0^x a_n t^n\, dt = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

et prouve le résultat. ■

EXERCICE 11-3.35 Montrer que, pour $|x| < 1$

$$\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$$

et montrer que cette égalité est encore valable pour $x = 1$.

11-3.2.5 Dérivation

Ici encore, nous commençons par l'étude de séries d'une variable réelle x.

THÉORÈME 11-3.36 Si $\sum a_k x^k$ est une série de rayon de convergence $R > 0$, la fonction

$$f :]-R, R[\to \mathbb{C} \quad f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

est de classe C^1 sur $]-R, R[$ et

$$\forall x \in]-R, R[\quad f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$$

Comme le rayon de convergence de la série dérivée est encore égal à R, on peut itérer le raisonnement. La fonction f est en fait de classe C^∞ sur $]-R, R[$, avec

$$\forall x \in]-R, R[\quad \forall p \in \mathbb{N}^*$$
$$f^{(p)}(x) = \sum_{n=p}^{+\infty} n(n-1)\cdots(n-p+1)\, a_n x^{n-p} = \sum_{n=0}^{+\infty} \frac{(n+p)!}{n!} a_{n+p} x^n$$

Démonstration : On pourrait faire intervenir le théorème de dérivation terme à terme, en invoquant une convergence uniforme locale de la série des dérivées. On peut aussi remarquer que la fonction continue définie sur $]-R, R[$ par

$$g(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$$

(série entière dont le rayon de convergence est aussi égal à R) est continue sur $]-R, R[$, avec

$$\int_0^x g(t)\, dt = \sum_{n=1}^{+\infty} a_n x^n = f(x) - a_0$$

ce qui prouve que f est C^1, avec $f' = g$. ■

REMARQUE 11-3.37 Le théorème d'intégration ou de dérivation s'applique sur $]-R, R[$. Si on veut dériver en $x = R$, ou intégrer sur $[0, R]$ une série entière particulière, il faudra justifier.

COROLLAIRE 11-3.38 L'expression de $f^{(p)}$ donne immédiatement

$$\forall p \in \mathbb{N} \qquad f^{(p)}(0) = p!\, a_p$$

ce qui peut s'écrire également

$$a_p = \frac{f^{(p)}(0)}{p!}$$

Ce résultat redémontre ce qui a déjà été vu (corollaire **11-3**.28), à savoir que, si une fonction f est, au voisinage de 0 (dans $\mathbb{R}$, *a fortiori* dans $\mathbb{C}$) égale à la somme d'une série entière, les coefficients de cette série sont uniques.

REMARQUE 11-3.39 On a donc là un moyen de "récupérer" les coefficients de la série entière en connaissant la fonction f : cette série s'écrit

$$\sum \frac{f^{(n)}(0)}{n!} x^n$$

série qu'on appelle **série de Taylor de** f en 0. Cette relation entre f et les coefficients (a_n) n'est pas toujours commode d'utilisation, car elle fait intervenir des dérivations. Comment par exemple majorer a_n si on a une majoration de f ? C'est l'objet de l'exercice qui suit :

EXERCICE 11-3.40 Dans le cas complexe, il est un moyen plus efficace de récupérer les coefficients de la série. La méthode suivie prendra tout son sens lorsqu'auront été étudiées les séries de Fourier : montrer que, si

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est somme d'une série entière dans le disque ouvert $D(0, R[$, alors

$$\forall n \in \mathbb{N} \quad \forall \rho \in [0, R[\quad a_n \rho^n = \frac{1}{2\pi} \int_0^{2\pi} f\left(\rho e^{i\theta}\right) e^{-in\theta}\, d\theta$$

EXERCICE 11-3.41 (Dérivation dans le cas complexe) : si Ω est un ouvert de $\mathbb{C}$, une application $f : \Omega \to \mathbb{C}$ est dite dérivable (comme fonction d'une variable complexe) en $z_0 \in \Omega$ si et seulement si la limite

$$L = \lim_{\substack{z \to z_0 \\ z \neq z_0}} \frac{f(z) - f(z_0)}{z - z_0}$$

existe dans $\mathbb{C}$. Comme dans le cas réel, L est noté $f'(z_0)$. Montrer que, si

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est somme d'une série entière de rayon de convergence $R > 0$, la fonction f est dérivable en tout point z_0 de $\Omega = D(0, R[$, avec

$$f'(z_0) = \sum_{n=1}^{+\infty} n a_n z_0^{n-1}$$

Il en résultera évidemment que f est indéfiniment dérivable sur Ω.

Indication : pour[12] $|h| < R - |z_0|$, $f(z_0 + h)$ est défini et

$$f(z_0 + h) - f(z_0) - h\sum_{n=1}^{+\infty} na_n\, z_0^{n-1} = \sum_{n=2}^{+\infty} a_n \left(\sum_{p=2}^{n} C_n^p\, h^p z_0^{n-p}\right)$$

par la formule du binôme. Si $|h| \leqslant \rho < R - |z_0|$, en déduire que

$$\left|f(z_0 + h) - f(z_0) - h\sum_{n=1}^{+\infty} na_n\, z_0^{n-1}\right| \leqslant \frac{|h|^2}{\rho^2}\sum_{n=2}^{+\infty} |a_n| \left(\sum_{p=2}^{n} C_n^p\, |\rho|^p\, |z_0|^{n-p}\right)$$

en prenant soin de vérifier la convergence de la série majorante. Conclure.

On en déduit par exemple que la fonction $z \mapsto \exp(z)$ est dérivable en tout point de $\mathbb{C}$ comme fonction de la variable complexe z et que cette fonction est égale à sa dérivée.

11-3.3 Développements en série entière

11-3.3.1 Définition

Ici encore les définitions sont données pour une fonction d'une variable réelle ou complexe :

DÉFINITION 11-3.42 *Soit Ω un ouvert de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une fonction f de Ω dans $\mathbb{C}$ est développable en série entière au voisinage[13] d'un point $z_0 \in \Omega$ si et seulement s'il existe une série entière $\sum a_k\, z^k$ de rayon de convergence $R > 0$ et un réel $\rho > 0$ (forcément $\leqslant R$) tel que*

$$\forall z \in \Omega \quad |z - z_0| < \rho \Rightarrow f(z) = \sum_{n=0}^{+\infty} a_n\, (z - z_0)^n$$

La fonction f est donc égale, dans un disque $D(z_0, \rho[$ centré en z_0, à la somme d'une série entière de la variable $Z = z - z_0$. On peut toujours se ramener en 0 en étudiant la fonction $f_{z_0}(h) = f(z_0 + h)$. Nous utiliserons par la suite l'abréviation

$$f \text{ est } \mathrm{DSE}_{z_0}$$

pour dire : "f est définie et développable en série entière au voisinage de z_0".

Les théorèmes qui précèdent montrent

PROPOSITION 11-3.43 Si f est DSE_{z_0}, il y a unicité de la série entière (de la variable $Z = z - z_0$) qui représente $f(z)$ au voisinage de z_0.

[12]Cette condition est peu contraignante si $R = +\infty$

[13]On dit aussi "en z_0" mais cette terminologie est moins précise.

11-3.3.2 Développements en séries entières et opérations

Les résultats de la section **11-3.1.5** et les théorème **11-3.**34 et **11-3.**36 donnent immédiatement :

PROPOSITION 11-3.44 Si f et g sont DSE$_{z_0}$, il en est de même de toute combinaison linéaire de f et g ainsi que de leur produit : si Ω est un ouvert de $\mathbb{R}$ ou $\mathbb{C}$, l'ensemble des fonctions $\Omega \to \mathbb{C}$ développables en série entière au voisinage d'un point $z_0 \in \Omega$ est une sous-algèbre de $\mathbb{C}^{\Omega}$.

PROPOSITION 11-3.45 Si I est un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{C}$ est une fonction continue qui est développable en série entière au voisinage de $x_0 \in I$, toute primitive de f dans I est DSE$_{x_0}$.

PROPOSITION 11-3.46 Si I est un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{C}$ est une fonction qui est développable en série entière au voisinage de $x_0 \in I$, alors f est de classe C^∞ au voisinage de x_0, et les dérivées successives de f sont toutes DSE$_{x_0}$.

EXERCICE 11-3.47 Soit

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n$$

la somme d'une série entière de rayon de convergence $R > 0$. Montrer que, si $a_0 \neq 0$, la fonction $g = \frac{1}{f}$ est définie au voisinage de zéro. Montrer que cette fonction est aussi développable en série entière au voisinage de 0.

Indication : on peut toujours supposer $a_0 = 1$ (ce qui simplifie un peu les calculs). Montrer qu'il existe une unique série entière $\sum b_k z^k$ dont le produit de Cauchy (formel) avec la série définissant f soit égal à 1. Montrer ensuite (en utilisant la caractérisation vue à l'exercice **11-3.**11) que le rayon de convergence de cette série n'est pas nul et conclure. Déduire de ce qui précède que, si f est DSE$_{z_0}$ vérifie $f(z_0) \neq 0$, alors $\frac{1}{f}$ est DSE$_{z_0}$.

11-3.3.3 Notion de fonction analytique

Il s'agit d'une notion commode lorsqu'on manipule des fonctions développables en séries entières :

DÉFINITION 11-3.48 *Si Ω est un ouvert de $\mathbb{R}$ ou $\mathbb{C}$, une fonction $f : \Omega \to \mathbb{C}$ est dite analytique sur Ω si et seulement si f est développable en série entière au voisinage de tout point de Ω :*

$$\forall z_0 \in \Omega \quad \exists (a_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}} \quad \exists \rho > 0 \quad |z - z_0| < \rho \Rightarrow f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$$

Il est à noter que la suite $(a_n)_{n\in\mathbb{N}}$ et ρ dépendent évidemment de z_0.

Nous noterons $\mathcal{H}(\Omega)$ l'ensemble des fonctions analytiques de Ω dans $\mathbb{C}$. C'est clairement une sous-algèbre de $\mathcal{C}^0(\Omega, \mathbb{C})$.

Un exemple fondamental de fonctions analytiques est donné par les fractions rationnelles :

PROPOSITION 11-3.49 Toute fraction rationnelle R est analytique sur son domaine de définition (c'est-à-dire $\mathbb{C}$ privé des pôles de R).

Démonstration : On utilise la décomposition de R en éléments simples. Soit $\{r_1, \ldots, r_m\}$ la liste des pôles de R. Comme la partie entière (polynomiale) de R est évidemment analytique, il suffit de montrer qu'il en est de même de toute fraction

$$Q_i(z) = \frac{1}{(z - r_i)^p}$$

où $p \leqslant m_i$, multiplicité du pôle r_i. On utilise le développement en série

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n \quad \text{pour } |z| < 1$$

qui donne, après $p-1$ dérivations

$$\frac{1}{(1-z)^p} = \sum_{n=0}^{+\infty} \frac{(n+p-1)!}{n!\,(p-1)!} z^n = \sum_{n=0}^{+\infty} \mathrm{C}_{n+p-1}^{p-1}\, z^n \quad \text{pour } |z| < 1$$

(à vrai dire, si on ne dispose pas du résultat de l'exercice **11-3**.41, ce développement est valable pour une variable réelle. Mais comme on sait, grâce au produit de Cauchy, que $z \mapsto \left[(1-z)^{-1}\right]^p$ est somme d'une série entière dans $D(0, 1[$ et qu'il y a unicité de cette série, le développement est aussi valable pour une variable complexe).

On écrira alors, pour z et $z_0 \in \mathbb{C} - \{r_1, \ldots, r_m\}$

$$\frac{1}{(z - r_i)^p} = \frac{1}{[(z - z_0) - (r_i - z_0)]^p} = \frac{(-1)^p}{(r_i - z_0)^p} \cdot \frac{1}{\left(1 - \dfrac{z - z_0}{r_i - z_0}\right)^p}$$

ce qui donnera, pour $|z - z_0| < |r_i - z_0|$

$$\frac{1}{(z - r_i)^p} = \frac{(-1)^p}{(r_i - z_0)^p} \sum_{n=0}^{+\infty} \mathrm{C}_{n+p-1}^{p-1} \left(\frac{z - z_0}{r_i - z_0}\right)^n$$

Chacun des éléments simples possédant donc un développement en série entière, on aura bien une écriture de la forme

$$R(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$$

valable pour $|z - z_0| < \inf\limits_{1 \leqslant i \leqslant m} |z_0 - r_i|$. ■

EXERCICE 11-3.50 Montrer que le rayon de convergence de cette série est exactement

$$R = \inf_i |z_0 - r_i|$$

qui représente la distance de z_0 à un des pôles dont il est le plus proche.

EXERCICE 11-3.51 En reprenant les calculs de l'exercice **11-3**.41 (et essentiellement un résultat de permutation de symboles de sommations), montrer que, si $\sum a_k z^k$ est une série entière de rayon de convergence $R > 0$, sa somme est une fonction analytique sur $D(0, R[$, et que

$$\forall z \in \mathbb{C} \quad |z - z_0| < R - |z_0| \Rightarrow f(z) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

Par exemple, la fonction $z \mapsto \exp(z)$ est analytique sur $\mathbb{C}$. L'équation fonctionnelle vérifiée par cette fonction nous donne d'ailleurs

$$\forall z \in \mathbb{C} \quad \exp(z) = \exp(z_0 + (z - z_0)) = \exp(z_0) \sum_{n=0}^{+\infty} \frac{(z - z_0)^n}{n!}$$

EXERCICE 11-3.52 (**Principe des zéros isolés**) Montrer que, si f est une fonction analytique dans un ouvert Ω connexe et s'il existe une suite $(z_n)_{n\in\mathbb{N}}$ non stationnaire de zéros de f (c'est-à-dire vérifiant $\forall n \quad f(z_n) = 0$) qui **converge vers un point** l **de** Ω, la fonction f est identiquement nulle (utiliser le développement en série entière de f en l pour montrer d'abord que f est identiquement nulle au voisinage de l).

11-3.3.4 Condition d'existence d'un D.S.E pour une fonction d'une variable réelle

Nous considérons dans cette section une fonction f d'une variable réelle définie sur un intervalle ouvert I contenant un point x_0. On cherche des conditions pour que f soit développable en série entière au voisinage de x_0. Le théorème **11-3**.36 et le corollaire **11-3**.38 donnent immédiatement :

PROPOSITION 11-3.53 Si f est développable en série entière au voisinage de x_0, alors f est C^∞ au voisinage de x_0 et la série de Taylor de f en x_0

$$\sum_{n=0}^{+\infty} \frac{f^{(n)}(x_0)}{n!} z^n$$

a un rayon de convergence non nul.

Démonstration : On sait en effet que si f est, dans un intervalle $]x_0 - \alpha, x_0 + \alpha[$, somme d'une série entière de la variable $(x - x_0)$, cette série s'écrit nécessairement

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \quad \blacksquare$$

ATTENTION ! LES CONDITIONS QUI PRÉCÈDENT NE SONT PAS SUFFISANTES POUR Q'UNE FONCTION DE CLASSE C^∞ SOIT DÉVELOPPABLE EN SÉRIE ENTIÈRE.

Par exemple, la fonction (qui nous a déjà servi de contre-exemple lorsque nous avons fait des rappels sur les développements limités)

$$f : \mathbb{R} \to \mathbb{R} \quad 0 \neq x \mapsto f(x) = e^{-\frac{1}{x^2}} \text{ prolongée en 0 par } f(0) = 0$$

est de classe C^∞ sur $\mathbb{R}$, et toutes ses dérivées à l'origine sont nulles. Il en résulte que le rayon de convergence de la série de Taylor à l'origine vaut $+\infty$ et cependant

$$\forall x \neq 0 \quad \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n = 0 \neq e^{-\frac{1}{x^2}} = f(x)$$

On pourrait bien sûr donner comme condition nécessaire et suffisante d'existence d'un développement en série entière en x_0 la propriété : "f est C^∞ au voisinage de x_0, la série de Taylor de f en x_0 a un rayon de convergence $R > 0$ et il existe $\alpha \in]0, R[$ tel que f soit dans $]x_0 - \alpha, x_0 + \alpha[$ somme de cette série de Taylor de la variable $(x - x_0)$" mais cette condition revient à dire que f est somme d'une série entière si et seulement si ... elle est somme d'une série entière ! Retenons cependant

PROPOSITION 11-3.54 Une fonction f d'une variable réelle définie au voisinage de x_0 est développable en série entière en x_0 si et seulement si f est C^∞ au voisinage de x_0 et

$$\exists \alpha > 0 \quad \forall x \in \mathbb{R} \quad |x - x_0| < \alpha \Rightarrow \lim_{n \to +\infty} \left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right| = 0$$

Nous verrons à l'exercice **11-3**.57 une condition nécessaire et suffisante portant sur des majorations de dérivées pour qu'une fonction soit dérivable en série entière (condition peu utilisable en pratique). Si on veut utiliser la proposition qui précède pour montrer qu'une fonction est développable en série entière, on pourra essayer de majorer la différence

$$\Delta_n(x) = \left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right|$$

à l'aide de l'inégalité de Taylor Lagrange, ou de la formule de Taylor avec reste intégral :

$$\begin{aligned} \Delta_n(x) &\leqslant \frac{|x - x_0|^{n+1}}{(n+1)!} \sup_{t \in [x_0, x]} \left| f^{(n+1)}(t) \right| \\ \Delta_n(x) &= \left| \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt \right| \end{aligned}$$

pour essayer de montrer que $\Delta_n(x)$ tend vers 0 lorsque $n \to +\infty$, pour x suffisamment proche de x_0.

EXERCICE 11-3.55 Montrer que, s'il existe un intervalle ouvert J centré en x_0 et une suite $(M_n)_{n\in\mathbb{N}}$ de réels positifs tels que

1) $\forall x \in J \quad \forall n \in \mathbb{N} \quad \left|f^{(n)}(x)\right| \leqslant M_n$
2) La série $\sum \frac{M_n}{n!} z^n$ a un rayon de convergence non nul

alors f est développable en série entière[14] en x_0.

EXEMPLE 11-3.56 SÉRIE DU BINÔME.

On considère la fonction $f : \left]-1,+\infty\right[\to \mathbb{R}^+$ définie par

$$f(x) = (1+x)^\alpha$$

où α est un réel arbitraire. On se propose de prouver que cette fonction est développable en série entière à l'origine. C'est évident si α est un entier naturel, puisque f est alors polynomiale. Si α est un entier négatif, on utilisera la dérivation de la série

$$\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad |x| < 1$$

comme on l'a fait lors de la démonstration de la proposition **11-3**.49. On supposera donc dans la suite $\alpha \notin \mathbb{Z}$. Comme

$$f^{(n)}(0) = \alpha(\alpha-1)\cdots(\alpha-n+1)$$

la série de Taylor de f à l'origine est donc

$$\sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

Ici, l'utilisation de la règle de D'Alembert pour obtenir le rayon de convergence de cette série est une bonne idée, et on obtient facilement $R=1$. On est alors amené à supposer $x \in \left]-1,1\right[$, et à étudier la différence entre $f(x)$ et les sommes partielles de la série de Taylor. Comme il n'est pas simple d'obtenir une majoration raisonnable de la dérivée d'ordre $n+1$ de la fonction f, on utilisera l'écriture de cette différence sous forme d'intégrale :

$$\Delta_n(x) = \left|\int_0^x \frac{\alpha(\alpha-1)\cdots(\alpha-n)}{n!}(1+t)^{\alpha-n-1}(x-t)^n \, dt\right|$$

Comme on vérifie que sur $[0,x]$ ou $[x,0]$ on a

$$\left|\frac{x-t}{1+t}\right| \leqslant |x|$$

[14] En utilisant ce résultat, on pourrait montrer que les fonctions exp, sin, cos sont analytiques sur $\mathbb{R}$. Comme ces fonctions ont été **définies** comme sommes de séries entières (à partir de la série de l'exponentielle), ce résultat est en fait conséquence des formules de trigonométrie, c'est à dire de l'équation fonctionnelle vérifiée par la fonction exponentielle.

on obtient la majoration

$$\forall x \in \left]-1,1\right[\quad \Delta_n(x) \leqslant \left| \frac{\alpha(\alpha-1)\cdots(\alpha-n)}{n!} |x|^n \right| \left| \int_0^x (1+t)^{\alpha-1}\, dt \right|$$

quantité qui tend vers zéro pour $n \to +\infty$ pour tout x de $]-1,1[$, puisque c'est le terme général d'une série convergente. On a donc bien

$$|x| < 1 \Rightarrow (1+x)^\alpha = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

EXERCICE 11-3.57 Soit f une fonction C^∞ définie sur un voisinage de 0. Montrer que f est développable en série entière en 0 si et seulement si

$$\exists \rho > 0 \quad \exists A > 0 \quad \exists C > 0 \quad \forall x \in \left]-\rho, \rho\right[\quad \forall n \in \mathbb{N} \quad \frac{\left|f^{(n)}(x)\right|}{n!} \leqslant C A^n$$

11-3.3.5 Méthodes de développements

Nous passons en revue différentes techniques pour obtenir des développements en séries entières, que nous appliquons à certaines fonctions "usuelles" :

Intégration et dérivation.

De la sommation de la série géométrique

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{et} \quad \frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad (\text{pour } |x| < 1)$$

nous avons déduit par dérivations

$$\forall p \in \mathbb{N} \quad \frac{1}{(1-x)^{p+1}} = \sum_{n=0}^{+\infty} C_{n+p}^p\, x^n \qquad (\text{pour } |x| < 1)$$

et par intégration

$$\ln(1-x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n} \quad \text{et} \quad \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} x^n \qquad (\text{pour } |x| < 1)$$

(on sait que cette dernière égalité est encore valable en $x = 1$). De même, à partir du développement de $\dfrac{1}{1+x^2}$ (série géométrique de raison $-x^2$, pour $|x| < 1$) on obtient :

$$\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)} x^{2n+1} \qquad (\text{pour } |x| < 1)$$

(avec ici encore égalité pour $x = \pm 1$).

EXERCICE 11-3.58 En utilisant la série du binôme, donner les développements en séries entière à l'origine de

$$\frac{1}{\sqrt{1-x^2}} \text{ et } \frac{1}{\sqrt{1+x^2}}$$

En déduire que

$$\arcsin x = \sum_{n=0}^{+\infty} \frac{(2n)!}{2^{2n}(n!)^2} \frac{x^{2n+1}}{2n+1} \qquad (\text{pour } |x| < 1)$$

$$\operatorname{argsh} x = \sum_{n=0}^{+\infty} \frac{(-1)^n (2n)!}{2^{2n}(n!)^2} \frac{x^{2n+1}}{2n+1} \qquad (\text{pour } |x| < 1)$$

et montrer que ces égalités sont encore valables en ± 1.

Opérations algébriques.

Partant du développement de la fonction exponentielle

$$e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \qquad (\text{pour } z \in \mathbb{R})$$

et, compte tenu des expressions des fonctions trigonométriques à l'aide de l'exponentielle, nous avons

$$\operatorname{sh} x = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!} \quad \text{et} \quad \operatorname{ch} x = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!} \qquad (\text{pour } x \in \mathbb{R})$$

$$\sin x = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \quad \text{et} \quad \cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n}}{(2n)!} \qquad (\text{pour } x \in \mathbb{R})$$

EXERCICE 11-3.59 Développer en série entière en 0 les fonctions

$$\ln\left(1 + \frac{x}{1+x^2}\right) \quad \text{et} \quad e^{2x}\cos x$$

et déterminer les domaines de validité des développements.

Utilisation d'une équation différentielle.

Nous avons obtenu la série du binôme par application de la formule de Taylor avec reste sous forme d'intégrale. Nous utilisons une équation différentielle pour retrouver ce résultat :

La fonction $f :]-1, +\infty[\to \mathbb{R}$ définie par $f(x) = (1+x)^\alpha$ est de classe C^∞ et vérifie

$$\forall x > -1 \quad f'(x) = \alpha(1+x)^{\alpha-1}$$

soit également

$$\forall x > -1 \quad (1+x) f'(x) - \alpha f(x) = 0$$

La fonction f vérifie donc l'équation différentielle linéaire homogène d'ordre 1

$$(1+x)\, y' - \alpha\, y = 0$$

avec la condition initiale $y(0) = 1$. On raisonne ensuite par analyse/synthèse :

- Recherche d'une série entière solution de cette équation : on suppose que la fonction $x \mapsto g(x) = \sum_{n=0}^{+\infty} a_n x^n$ (somme d'une série dont le rayon de convergence R est supposé > 0) vérifie l'équation différentielle au voisinage de zéro. On a alors

$$\forall x \in]-R, R[\quad (1+x) g'(x) - \alpha g(x) = (1+x) \sum_{n=1}^{+\infty} n a_n x^{n-1} - \alpha \sum_{n=0}^{+\infty} a_n x^n$$

soit finalement

$$\forall x \in]-R, R[\quad (1+x) g'(x) - \alpha g(x) = \sum_{n=0}^{+\infty} ((n+1) a_{n+1} + (n-\alpha) a_n) x^n$$

Cette fonction est identiquement nulle sur un voisinage de 0. Tous les coefficients de la série sont donc nécessairement nuls (unicité d'un développement en série entière), soit

$$\forall n \geqslant 0 \quad a_{n+1} = \frac{\alpha - n}{n+1} a_n$$

Ceci équivaut à

$$\forall n \geqslant 1 \quad a_n = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} a_0$$

qui est donc la condition nécessaire et suffisante pour que g vérifie l'équation différentielle (en fait sur $]-R, R[$). Si de plus g vérifie la même condition initiale que f, on a $a_0 = 1$ et donc

$$\forall x \in]-R, R[\quad g(x) = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

- Réciproquement, si $\alpha \notin \mathbb{N}$, le rayon de convergence de cette série est 1 (D'Alembert). Et, comme les coefficients de la série vérifient la relation de récurrence étudiée précédemment, la fonction g vérifie effectivement l'équation différentielle sur $]-1, 1[$, avec la condition initiale $g(0) = 1$.

- Pour identifier f et g, on fait ensuite appel à la théorie des équations différentielles : un théorème (dû à Cauchy) affirme que l'équation considérée possède sur $]-1, 1[$ (intervalle sur lequel le coefficient de y' ne s'annule pas) une unique solution vérifiant la condition initiale $f(0) = 1$. On a donc bien

$$\forall x \in]-1, 1[\quad f(x) = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

La méthode a bien fonctionné car l'équation différentielle était "simple" : linéaire à coefficients polynomiaux, ce qui amène à des calculs raisonnables pour obtenir les coefficients a_n. Avec d'autres équations, les calculs pourraient être inextricables.

EXERCICE 11-3.60 Développer en série entière à l'origine la fonction

$$f(x) = e^{x^2} \int_x^{+\infty} e^{-t^2} dt$$

(on admet que $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$).

A l'inverse, on peut utiliser une équation différentielle pour sommer certaines séries entières, ou des séries entières pour résoudre certaines équations différentielles :

EXEMPLE 11-3.61 Le rayon de convergence de la série

$$\sum_{n=0}^{+\infty} \frac{2^{2n}(n!)^2}{(2n+1)!} x^{2n}$$

se calcule par la règle de D'Alembert. Si on note a_n le coefficient de x^{2n}, on a en effet

$$\frac{a_{n+1}}{a_n} = \frac{2n+2}{2n+3}$$

En écrivant cette relation sous la forme

$$(2n+3)\, a_{n+1} - (2n+2)\, a_n = 0$$

soit finalement, pour $|x| < 1$

$$\sum_{n=0}^{+\infty} [(2n+3)\, a_{n+1} - (2n+2)\, a_n]\, x^{2n} = 0$$

on montre que la somme de la série est, dans $]-1, 1[$, solution de l'équation différentielle

$$\left(x - x^3\right) y' + \left(1 - 2x^2\right) y = 1$$

ce qui permet de prouver que, sur $]-1, 1[$

$$\sum_{n=0}^{+\infty} \frac{2^{2n}(n!)^2}{(2n+1)!} x^{2n} = \frac{\arcsin x}{x\sqrt{1-x^2}}$$

EXERCICE 11-3.62 Résoudre l'équation différentielle

$$xy'' - y' + 4x^3 y = 0$$

Permutation de symboles de sommation.

EXERCICE 11-3.63 Développer en série entière la fonction

$$f(x) = \int_0^1 e^{ixt^5} dt$$

EXERCICE 11-3.64 Montrer que, pour $|a| < 1$, la formule

$$f(x) = \sum_{n=0}^{+\infty} \sin(a^n x)$$

définit une fonction indéfiniment dérivable sur $\mathbb{R}$, et montrer que f est développable en série entière en 0.

EXERCICE 11-3.65 Fonction de Bessel J_0. On définit

$$J_0(x) = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \cos(x \sin t)\ dt$$

Montrer que J_0 est développable en série entière à l'origine. Retrouver ce résultat en montrant que J_0 est solution de l'équation différentielle

$$x\, y'' + y' + x\, y = 0$$

11-3.4 Exemples d'utilisations des séries entières

Outre l'intégration de certaines équations différentielles développée à la section précédente, nous donnons ici quelques exemples classiques où les séries entières s'avèrent très utiles.

11-3.4.1 Sommation de certaines séries

EXEMPLE 11-3.66 Calculer la somme de la série

$$S = \sum_{n=0}^{+\infty} \frac{1}{(6n+2)(6n+5)}$$

Comme

$$\frac{1}{(6n+2)(6n+5)} = \frac{1}{3}\left[\frac{1}{6n+2} - \frac{1}{6n+5}\right]$$

on a aussi

$$S = \sum_{n=0}^{+\infty} \frac{(-1)^n}{3n+2}$$

Il est alors naturel de considérer la fonction

$$S(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{3n+2} x^{3n+2}$$

qui est définie sur $]-1, 1]$ et vérifie, dans l'intervalle **ouvert** $]-1, 1[$

$$S'(x) = \sum_{n=0}^{+\infty} (-1)^n x^{3n+1} = \frac{x}{1+x^3}$$

on en déduit que, pour $|x| < 1$

$$S(x) = \int_0^x \frac{t}{1+t^3} dt$$

et il reste à montrer que cette égalité est encore valable pour $x = 1$.

Plus généralement, si on connaît le rayon de convergence R et la somme

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n$$

d'une série entière, on pourra calculer la somme

$$T(x) = \sum_{n=0}^{+\infty} P(n) a_n x^n$$

si P est un polynôme (on sait que cette série a même rayon de convergence que la série de départ). Il suffira d'exprimer le polynôme P dans la base

$$1, X, X(X-1), X(X-1)(X-2), \ldots$$

pour obtenir une expression de T à l'aide de S et de ses dérivées successives.

EXERCICE 11-3.67 Calculer

$$\sum_{n=0}^{+\infty} \frac{(-1)^n n^2}{(2n+1)!} x^{2n}$$

De même, pour calculer

$$U(x) = \sum_{n=0}^{+\infty} R(n) a_n x^n$$

où R est une fraction rationnelle, on pourra décomposer R en éléments simples. On se ramène alors à calculer des sommes de séries de la forme

$$V(x) = \sum_{n=0}^{+\infty} \frac{a_n}{(n+\alpha)^k} x^n$$

Pour essayer d'expliciter V, on pourra dériver la fonction

$$x^\alpha V(x) = \sum_{n=0}^{+\infty} \frac{a_n}{(n+\alpha)^k} x^{n+\alpha}$$

ce qui diminuera la valeur de k.

EXERCICE 11-3.68 Calculer

$$\sum_{n=1}^{+\infty} \frac{n^2+4n-1}{n+4} \frac{x^n}{n!}$$

EXERCICE 11-3.69 Calcul de la somme de la série

$$\sum_{n=1}^{+\infty} \frac{e^{inx}}{n}$$

pour $x \in]0, 2\pi[$ (voir remarque **11-2.30**). Indication : calculer, pour $r \in]0, 1[$ la somme

$$\sum_{n=1}^{+\infty} \frac{e^{inx}}{n} r^n$$

et conclure, en utilisant par exemple l'exercice **11-3.33**.

11-3.4.2 Régularité de certaines fonctions

La somme d'une série entière est, dans l'intervalle ouvert de convergence, de classe $\mathcal{C}^\infty$. C'est un résultat qui permet parfois de prouver très rapidement la régularité d'une fonction :

EXERCICE 11-3.70 Montrer que la fonction

$$f(x) = \frac{\ln(1+x).\sin(x)}{1-\cos x}$$

est prolongeable en une fonction continue à l'origine, et que cette fonction est $\mathcal{C}^\infty$ au voisinage de 0.

EXERCICE 11-3.71 Montrer que la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(t) = \mathrm{ch}\sqrt{t}$ pour $t \geqslant 0$ et $f(t) = \cos\sqrt{-t}$ pour $t < 0$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

11-3.4.3 Calculs approchés

EXERCICE 11-3.72 Démontrer que

$$\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$$

et, utilisant le développement en série de la fonction arctan, donner une valeur approchée à 10^{-6} près de π.

EXERCICE 11-3.73 Montrer que

$$\int_0^1 x^{-x}\,dx = \sum_{n=1}^{+\infty}\frac{1}{n^n}$$

et donner une valeur approchée de cette intégrale à 10^{-10} près.

11-3.4.4 Etude de suites, analyse combinatoire

EXERCICE 11-3.74 Nombre manières de payer une somme de n francs avec des pièces de 1, 2 et 5 francs. Il s'agit du nombre de solutions en entiers de l'équation

$$p + 2q + 5r = n$$

Montrer qu'il s'agit du coefficient de x^n dans le développement en série entière de

$$R(x) = \frac{1}{(1-x)(1-x^2)(1-x^5)}$$

EXERCICE 11-3.75 Même exercice en euros.

EXERCICE 11-3.76 On considère une suite $(a_n)_{n\in\mathbb{N}}$ de réels définie par

$$a_0 = 1 \quad \text{et} \quad \forall n \geqslant 0 \quad a_{n+1} = \sum_{k=0}^{n} a_k\, a_{n-k}$$

Expliciter a_n en fonction de n, en faisant intervenir la "série génératrice" de la suite (a_n), c'est-à-dire la série entière

$$\sum_{n=0}^{+\infty} a_n\, x^n$$

EXERCICE 11-3.77 On appelle D_n le nombre de partitions de l'ensemble $\{1, \dots, n\}$ (avec, par convention, $D_0 = 1$). Montrer que

$$\forall n \geqslant 0 \quad D_{n+1} = \sum_{k=0}^{n} \mathrm{C}_n^k D_k$$

En utilisant la fonction

$$f(x) = \sum_{n=0}^{+\infty} \frac{D_n}{n!} x^n$$

(on commencera par montrer que le rayon de convergence est non nul), montrer que

$$D_n = \frac{1}{e} \sum_{p=0}^{+\infty} \frac{p^n}{p!}$$

11-3.4.5 Séries entières dans un e.v.n complet, dans une algèbre de Banach

Nous nous sommes limités dans ce qui précède aux séries entières pour lesquelles la variable z était réelle ou complexe, et la suite (a_n) des coefficients était une suite complexe. Deux généralisations sont possibles :

- Si $\mathbb{E}$ est un $\mathbb{K}$-ev normé complet, et $(a_n)_{n \in \mathbb{N}}$ est une suite de vecteurs de $\mathbb{E}$, on peut considérer des séries de fonctions d'une variable $z \in \mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$)

  $$\sum_{n=0}^{+\infty} z^n \, a_n$$

  (on met cette fois le scalaire devant le vecteur). Beaucoup des résultats étudiés précédemment subsistent, les démonstrations étant analogues : rayon de convergence, continuité, dérivabilité, intégration ....

- Si $\mathbb{E}$ est une $\mathbb{K}$-algèbre de Banach, on peut cette fois considérer des séries de la forme

  $$\sum_{n=0}^{+\infty} a_n \, X^n$$

  où $(a_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ est une suite réelle ou complexe et X est une variable dans $\mathbb{E}$. Si R est le rayon de convergence de la série entière

  $$\sum_{n=0}^{+\infty} a_n \, z^n$$

  pour $X \in \mathbb{E}$ vérifiant $\|X\| < R$, on a la majoration

  $$\|a_n \, X^n\| \leqslant |a_n| \, \|X\|^n$$

qui prouve la convergence absolue de la série $\sum a_n X^n$. La même majoration prouve d'ailleurs qu'il y a convergence normale sur toute boule $B(0_{\mathbb{E}}, \rho]$ pour $\rho < R$, et montre que la somme de la série

$$S : B(0_{\mathbb{E}}, R[\to \mathbb{E} \qquad X \mapsto S(X) = \sum_{n=0}^{+\infty} a_n X^n$$

est une fonction continue.

Nous avons déjà rencontré à deux reprises ce genre de séries :

$$\exp(X) = \sum_{n=0}^{+\infty} \frac{X^n}{n!} \quad \text{et} \quad (1_{\mathbb{E}} - X)^{-1} = \sum_{n=0}^{+\infty} X^n$$

la première série convergeant en tout point de $\mathbb{E}$, la seconde[15] dans $B(0_{\mathbb{E}}, 1[$. L'exercice qui suit montre que d'autres séries rencontrées dans le cas réel ou complexe peuvent être utilisées dans une algèbre de Banach :

EXERCICE 11-3.78 Montrer que, dans une algèbre de Banach,

$$l(X) = -\sum_{n=1}^{+\infty} \frac{X^n}{n}$$

définit une fonction continue sur la boule unité ouverte. Montrer[16] que

$$\|X\| < 1 \Rightarrow \exp[l(X)] = 1_{\mathbb{E}} - X$$

Indication : le résultat est assuré dans $\mathbb{R}$. Dans le cas général, on pourrait raisonner par interversion des sommations. Il est plus simple de fixer X de norme < 1, de considérer l'application

$$\varphi : [0,1] \to \mathbb{E} \qquad t \mapsto \varphi(t) = (1_{\mathbb{E}} - tX) \times \exp[-l(tX)]$$

et de démontrer que cette fonction est dérivable, de dérivée nulle.

EXERCICE 11-3.79 Utiliser ce qui précède pour essayer de montrer que

$$\exp : \mathcal{M}_n(\mathbb{C}) \to \mathcal{GL}_n(\mathbb{C})$$

est une surjection.

[15] C'est grâce à cette série que nous avons montré que l'ensemble des éléments inversibles de $\mathbb{E}$ est un ouvert

[16] On a donc défini un logarithme dans $B(1_{\mathbb{E}}, 1[$ par

$$\ln(A) = -\sum_{n=1}^{+\infty} \frac{(1_{\mathbb{E}} - A)^n}{n}$$

11-4 Exercices

EXERCICE 11-4.1 Soit f continue $\mathbb{R} \to \mathbb{R}$ avec $\forall x \neq 0 \;\; |f(x)| < |x|$. Convergence simple et uniforme de la suite f_n définie par $f_0(x) = f(x)$ et $\forall n \geqslant 1 \;\; f_n(x) = f(f_{n-1}(x))$. Etudier en particulier les cas $f(x) = \sin x$ et $f(x) = |x|e^{-\frac{1}{|x|}}$.

EXERCICE 11-4.2 Montrer que, pour $x \in [0,1]$, les relations $P_0(x) = 1$ et

$$P_n(x) = 1 + \int_0^x P_{n-1}(t - t^2)dt$$

définissent une suite de polynômes. Que peut-on dire de $P_n(x) + P_n(1-x)$? Montrer que la suite P_n converge uniformément sur [0,1] vers une fonction f dérivable vérifiant

$$\forall x \in [0,1] \;\; f'(x) = f(x - x^2)$$

EXERCICE 11-4.3

- Montrer que $\sum\limits_{n \geqslant 1}(-1)^{n+1} \ln(1 + \frac{t}{n})$ converge simplement pour $t \in]-1, +\infty[$ et que la fonction ainsi définie est C^∞.
- Montrer que $g(x) = \sum\limits_{n=1}^{\infty} x^n \frac{\sin nx}{n}$ est C^1 sur]-1,1[et calculer g'.
- Calculer $\lim\limits_{x \to +\infty} \sum\limits_{n>0} (-1)^n \ln\left(1 + \frac{x^2}{n(1+x^2)}\right)$

EXERCICE 11-4.4 Théorème de Dini. Soit X un compact d'un espace normé et $(f_n) : X \to \mathbb{R}$ une suite croissante $(f_n \leqslant f_{n+1})$ de fonctions continues sur X convergeant simplement vers une fonction continue f. Montrer que la convergence est uniforme. Donner un contre-exemple lorsque X n'est pas compact.

EXERCICE 11-4.5 On pose $u_n(x) = (-1)^{n-1} \frac{n}{n^2 + x^2}$. Déterminer le domaine de définition de la fonction $U(x) = \sum\limits_{n=1}^{+\infty} u_n(x)$. Calculer $U(0)$. Déterminer $\lim\limits_{x \to +\infty} U(x)$. U est-elle C^∞ sur son domaine de définition ?

EXERCICE 11-4.6 Soit $\sum a_n$ une série à termes strictement positifs convergente. On pose

$$f(x) = \sum_{n=0}^{+\infty} a_n e^{-nx}$$

Montrer que f et $\ln f$ sont convexes sur $\mathbb{R}^+$.

EXERCICE 11-4.7 Soit $D = \{z \in \mathbb{C} \mid |z| < 1\}$.

1. Trouver toutes les $R \in \mathbb{C}(X)$ telles que $\forall z \in D \;\; R(z^2) - R(z) = \frac{z}{z^2 - 1}$.
2. Trouver toutes les $f \in C^0(D, \mathbb{C})$ telles que $\forall z \in D \;\; f(z^2) - f(z) = \frac{z}{z^2 - 1}$.
3. Déterminer le domaine de convergence et la valeur de $\sum\limits_{n=1}^{+\infty} \frac{z^{2^{n-1}}}{1 - z^{2^n}}$.

EXERCICE 11-4.8 Convergence de $f(t) = \sum_{n\geqslant 0} e^{-t\sqrt{n}}$. Calculer la dérivée de f et tracer son graphe. Trouver un équivalent de f en 0.

EXERCICE 11-4.9 Soit $f \in \mathcal{C}^0([0,1],\mathbb{R})$ et $u_n(x) = \prod_{k=1}^{n} \left(1 + \frac{x}{n} f\left(\frac{k}{n}\right)\right)$.

1. Etudier la convergence simple et déterminer la limite u de la suite u_n.
2. La convergence est-elle uniforme ?
3. On suppose f de classe C^1. Déterminer un équivalent, pour $n \to +\infty$, de la différence $u(x) - u_n(x)$.

EXERCICE 11-4.10 Calculer $\lim_{n\to+\infty} \sum_{k=1}^{n-1} \left(\frac{k}{n}\right)^n$

EXERCICE 11-4.11 Soit $\alpha \in \mathbb{N}$ et $t \in [0,1[$. Montrer que la série $\sum_{n=1}^{+\infty} \frac{n^\alpha t^n}{1-t^n}$ converge et donner un équivalent de sa somme pour $t \to 1^-$.

EXERCICE 11-4.12 Donner un équivalent, quand $x \to 0$, de

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^3} \arctan(nx) \quad \text{et de} \quad g(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \arctan(nx)$$

EXERCICE 11-4.13 Soit $P \in \mathbb{R}[X]$ tel que $\forall x \geqslant 0\ P(x) \leqslant e^x$. Montrer que

$$\exists M \in \mathbb{N}^* \quad \forall x > 0 \quad P(x) \leqslant \sum_{k=0}^{M} \frac{x^k}{k!}$$

EXERCICE 11-4.14 Rayon de convergence des séries :

$$\sum_{n\geqslant 0} \frac{z^n}{2 - \sin n\alpha} \quad (\alpha \in \mathbb{R}) \qquad \sum_n a^{n^2} z^{1+2+\cdots+n} \quad (a > 0)$$

EXERCICE 11-4.15 Soit $\sum a_n z^n$ une série entière de rayon de convergence $R > 0$. On note $S_n = \sum_{k=0}^{n} a_k$. Que peut-on dire du rayon de convergence de la série $\sum_n S_n x^n$?

EXERCICE 11-4.16 Montrer que la fonction $g(x) = \frac{1}{\sin x} - \frac{1}{x}$ est prolongeable en une fonction $\mathcal{C}^\infty$ sur $]-\pi, \pi[$.

EXERCICE 11-4.17 Soit U un ouvert connexe de $\mathbb{C}$, et $f : U \to \mathbb{C}$ analytique dans U. Pour $a \in U$ et r assez petit, évaluer, à l'aide du développement en série entière de f en a, l'intégrale

$$\frac{1}{2\pi} \int_0^{2\pi} |f(a + re^{i\theta})|^2 d\theta$$

En déduire que, si $|f|$ présente un maximum local en un point $a \in U$, alors f est constante au voisinage de a. Que peut-on en déduire sur U ?

EXERCICE 11-4.18 Inégalités de Cauchy.
Soit $\sum a_n z^n$ une série entière complexe de rayon de convergence $R > 0$ et de somme f. Pour $r \in [0, R[$ on pose $M(r) = \sup\{|f(z)| \mid |z| = r\}$. Montrer que

$$\forall n \in \mathbb{N} \quad |a_n| r^n \leqslant M(r)$$

En déduire que, pour $R = +\infty$, si f est bornée sur $\mathbb{C}$ alors f est constante.

EXERCICE 11-4.19 Soit $(a_n)_{n\in\mathbb{N}}$ une suite complexe convergeant vers a. Quel est le rayon de convergence de la série $\sum_{n\geqslant 0} \frac{a_n}{n!} x^n$? Si f désigne la somme de cette série, déterminer $\lim_{x\to+\infty} e^{-x} f(x)$.

EXERCICE 11-4.20 Soit a_n une suite réelle positive bornée, avec $\sum a_n$ divergente.

1. Déterminer $\lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n x^n$.
2. Si b_n est une suite réelle avec $b_n \sim a_n$ $(n \to +\infty)$, quel est le rayon de convergence de la série $\sum_{n=0}^{+\infty} b_n x^n$? Déterminer

$$\lim_{x\to 1^-} \frac{\sum_n b_n x^n}{\sum_n a_n x^n}$$

Application : trouver un équivalent en 1^- de $\sum_{n\geqslant 1} n^{\alpha-1} x^n$ pour $\alpha > 0$.

EXERCICE 11-4.21 Théorème de Tauber

1. Soit a_n une suite réelle telle que $\lim_{n\to+\infty} \frac{a_1 + 2a_2 + \cdots + n a_n}{n} = 0$. Montrer que $\sum a_n z^n$ a un rayon de convergence supérieur ou égal à 1.
2. On suppose de plus $a_n = o\left(\frac{1}{n}\right)$. Pour $r \in [0, 1[$ on pose $f(r) = \sum_{n\geqslant 0} a_n r^n$. On suppose que $\lim_{r\to 1^-} f(r) = l$ existe. Montrer que $\sum_{n\geqslant 0} a_n = l$ (on pourra étudier $f\left(1 - \frac{1}{n}\right) - \sum_{k=0}^{n} a_k$).

EXERCICE 11-4.22 Logarithme complexe
Soit $\mathcal{D} = \mathbb{C} - \mathbb{R}$. Pour $z \in \mathcal{D}$ on pose $\varphi(z) = \int_0^1 \frac{z-1}{tz + 1 - t} dt$. Montrer que φ est continue sur $\mathcal{D}$, vérifie $\varphi(1) = 0$ et $e^{\varphi(z)} = z$. Montrer que φ est la seule fonction continue sur $\mathcal{D}$ vérifiant ces propriétés. Développer la fonction φ en série entière au voisinage de 1.

EXERCICE 11-4.23 Soit $a_n = \int_0^1 \left(\frac{1+t^2}{2}\right)^n dt$ pour $n \in \mathbb{N}$. Montrer que la suite (a_n) tend vers 0, et que la série $\sum (-1)^n a_n$ converge. Calculer sa somme. Pour quelles valeurs de $x \in \mathbb{R}$ la série $\sum a_n x^n$ est-elle convergente ? Calculer sa somme. Donner un équivalent de a_n pour $n \to +\infty$.

EXERCICE 11-4.24 Etude de la série de terme général $\frac{n!}{(n+1)(n+2)\cdots(2n+1)}x^n$. Calcul de la somme.

EXERCICE 11-4.25 Soit $f : \mathbb{R} \to \mathbb{R}$ continue, telle que

$$\exists q \in]-1,1[\quad \forall x \in \mathbb{R} \quad f(x) = (1-qx)f(qx)$$

Montrer que f est développable en série entière sur $\mathbb{R}$.

EXERCICE 11-4.26 Soit f développable en série entière sur $]-R,R[$: $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. On considère l'équation différentielle

$$(\mathcal{E}) \begin{cases} y' - y = f\left(\frac{1}{x}\right) \\ x > \frac{1}{R} \end{cases}$$

- Montrer que $(\mathcal{E})$ possède une seule solution bornée au voisinage de $+\infty$. On la note y.
- Soit $\psi_n(x) = -\int_x^{+\infty} \frac{e^{-t}}{t^n} dt$. Montrer que $y(x) = e^{-x} \sum_{n=0}^{+\infty} a_n \psi_n(x)$.
- Donner un équivalent de ψ_n en $+\infty$.
- Montrer que y est équivalent à $-f\left(\frac{1}{x}\right)$ en $+\infty$.

EXERCICE 11-4.27 Montrer que $f(x) = \int_0^{\pi} \cos(t - x\sin t)dt$ est aussi égale à

$$\frac{1}{2}\int_{-\pi}^{\pi} e^{i(t - x\sin t)}dt$$

En donner le développement en série entière.

Chapitre 12

Intégration sur un intervalle quelconque

Rappel : si I est un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{C}$, on dit que f est continue par morceaux sur I si et seulement si, pour tout segment $K \subset I$, la fonction $f|_K$ est continue par morceaux. L'ensemble de ces fonctions forme un $\mathbb{C}$-ev noté $\mathcal{C}_m(I, \mathbb{C})$. On considèrera aussi l'espace réel $\mathcal{C}_m(I, \mathbb{R})$ et le sous-ensemble (stable par addition et homothétie de rapport positif) $\mathcal{C}_m^+(I)$ formé des fonctions à valeurs réelles positives.

12-1 Fonctions intégrables positives

12-1.1 Définition et propriétés élémentaires

12-1.1.1 Définition

DÉFINITION 12-1.1 (Fonction intégrable positive) *Si $f \in \mathcal{C}_m^+(I)$, on dit que f est intégrable (ou sommable) sur I si et seulement si l'ensemble*

$$\mathcal{A} = \left\{ \int_K f \text{ , } K \text{ segment inclus dans } I \right\}$$

*est un sous-ensemble majoré de $\mathbb{R}^+$. On définit alors **l'intégrale de f sur I** comme étant le réel positif*

$$\int_I f \underset{déf}{=} \sup_{\substack{K \text{ segment} \\ K \subset I}} \int_K f$$

On "mesure" en fait l'aire de la portion du plan déterminée par

$$\{(x,y) \in \mathbb{R}^2 \mid x \in I \text{ et } 0 \leqslant y \leqslant f(x)\}$$

que l'on "approche", par l'intérieur, par des sous ensembles de la forme

$$\{(x,y) \in \mathbb{R}^2 \mid x \in K \text{ et } 0 \leqslant y \leqslant f(x)\}$$

avec K segment inclus dans I, dont les aires valent $\int_K f$.

EXEMPLE 12-1.2 La fonction $f : t \mapsto \frac{1}{t^2}$ est intégrable sur $[1, +\infty[$ et son intégrale vaut $\int_{[1,+\infty[} f = 1$. Sur $]0, 1]$ cette fonction n'est pas intégrable.

DÉFINITION 12-1.3 *Si $f \in C_m^+(I)$ et si J est un sous-intervalle de I, on dit que f est intégrable sur J ssi $f|_J$ l'est et, par abus (commode) d'écriture, on note*

$$\int_J f = \int_J f|_J$$

EXEMPLE 12-1.4 Si $f : [a, b] \to \mathbb{C}$ est continue par morceaux, elle est intégrable sur $[a, b]$, $]a, b]$, $[a, b[$ et $]a, b[$ et si J désigne un de ces intervalles on a

$$\int_J f = \int_a^b f(t)dt$$

Ceci est une conséquence facile de la continuité d'une intégrale fonction d'une de ses bornes.

12-1.1.2 Propriétés élémentaires

- **Homogénéité** : si $f \in C_m^+(I)$ est intégrable sur I et si $\lambda \geqslant 0$ alors λf est intégrable et $\int_I \lambda f = \lambda \int_I f$. C'est une conséquence immédiate de la définition.

- **Additivité** : si f et $g \in C_m^+(I)$ sont intégrables sur I, il en est de même de $f + g$ et

$$\int_I (f + g) = \int_I f + \int_I g$$

Démonstration : Si K est un segment inclus dans I, on a

$$\int_K (f + g) = \int_K f + \int_K g \leqslant \int_I f + \int_I g$$

ce qui montre que $f+g$ est intégrable sur I et vérifie $\int_I (f + g) \leqslant \int_I f + \int_I g$. Pour obtenir l'inégalité inverse, on se donne un réel $\varepsilon > 0$ arbitraire. Il existe alors des segments K_1 et $K_2 \subset I$ tels que

$$\int_{K_1} f > \int_I f - \frac{\varepsilon}{2} \quad \text{et} \quad \int_{K_2} g > \int_I g - \frac{\varepsilon}{2}$$

Si K est le plus petit segment inclus dans I contenant $K_1 \cup K_2$ on a alors

$$\int_K (f+g) \geqslant \int_{K_1} f + \int_{K_2} g > \int_I f + \int_I g - \varepsilon$$

ce qui démontre le résultat. ∎

- **Croissance, domination** : si f et $g \in \mathcal{C}_m^+(I)$ vérifient $0 \leqslant f \leqslant g$ alors

$$g \text{ intégrable sur } I \Rightarrow f \text{ intégrable sur } I$$

donc bien sûr, si f n'est pas intégrable, il en est de même de g. Dans le cas de l'intégrabilité on a

$$0 \leqslant f \leqslant g \Rightarrow \int_I f \leqslant \int_I g$$

- **Positivité** : par construction, si $f \in \mathcal{C}_m^+(I)$ est intégrable, son intégrale est positive. On a d'ailleurs stricte positivité en général :

$$f \in \mathcal{C}_m^+(I) \text{ et } \int_I f = 0 \Longrightarrow f \text{ est nulle en tout point de continuité}$$

puisqu'en effet si x_0 est un point de continuité de f et si K est un segment inclus dans I contenant x_0, on a $0 \leqslant \int_K f \leqslant \int_I f = 0$ et on sait que ceci entraîne la nullité de f en x_0. Si f est d'intégrale nulle, sur chaque segment il n'y a qu'un nombre fini de points où f prend une valeur non nulle.

12-1.2 Caractérisation à l'aide d'une suite croissante de segments

Tout intervalle de $\mathbb{R}$ peut s'écrire

$$I = \bigcup_{n \in \mathbb{N}} K_n$$

où $(K_n)_{n \in \mathbb{N}}$ est une suite de segments croissante pour l'inclusion ($K_n \subset K_{n+1}$). Si $\alpha = \inf I \in \mathbb{R} \cup \{-\infty\}$ et $\beta = \sup I \in \mathbb{R} \cup \{+\infty\}$ et $K_n = [\alpha_n, \beta_n]$ est une telle suite, la suite α_n est décroissante et tend vers α, la suite β_n est croissante et tend vers β. Si de plus $\alpha \in I$ (par exemple), la suite α_n est stationnaire. On utilisera parfois en abrégé la notation

$$I = \bigcup_{\nearrow} K_n$$

pour "l'intervalle I est réunion de la suite croissante de segments K_n".

THÉORÈME 12-1.5 Soit $f \in \mathcal{C}_m^+(I)$. S'il existe une suite croissante (pour l'inclusion) de segments K_n dont la réunion est égale à I et telle que la suite des intégrales

$$\left(\int_{K_n} f\right)_{n \in \mathbb{N}}$$

soit majorée, alors la fonction f est intégrable sur I et

$$\int_I f = \sup_{n\in\mathbb{N}} \int_{K_n} f = \lim_{n\to+\infty} \int_{K_n} f$$

Réciproquement, si f est intégrable sur I, l'égalité précédente est valable pour toute suite croissante de segments d'union égale à I.

Démonstration : Comme $f \geqslant 0$, la suite des intégrales considérée est croissante. Si elle est majorée, elle converge vers sa borne supérieure. Si $K = [a, b] \subset I$, il existe un entier n_1 tel que $a \in K_{n_1}$ et un entier n_2 tel que $b \in K_{n_2}$. On a donc, pour $m \geqslant \max(n_1, n_2)$, $K \subset K_m$. Comme f est positive on a, pour ces valeurs de m,

$$\int_K f \leqslant \int_{K_m} f \leqslant \lim_{n\to+\infty} \int_{K_n} f$$

f est donc intégrable sur I et vérifie

$$\int_I f \leqslant \lim_{n\to+\infty} \int_{K_n} f$$

L'inégalité inverse est une évidence, puisque $\int_I f$ majore par définition la suite $\left(\int_{K_n} f\right)_{n\in\mathbb{N}}$. Enfin si réciproquement f est supposée intégrable, le raisonnement précédent est valable pour toute suite croissante de segments de réunion I. ■

On notera l'analogie de ce résultat avec celui obtenu pour caractériser les familles sommables de réels positifs. Le parallélisme n'est pas fortuit, puisqu'une théorie "moderne" de l'intégration (dite théorie de la mesure) englobe ces différentes constructions.

12-1.3 Additivité par rapport à l'intervalle d'intégration

THÉORÈME 12-1.6 Si $I = I_1 \cup I_2$ est une partition de I en deux intervalles non vides et si $f : I \to \mathbb{R}^+$ est continue par morceaux, f est intégrable sur I si et seulement si elle l'est sur I_1 et I_2 et on a alors

$$\int_I f = \int_{I_1} f + \int_{I_2} f$$

Démonstration : Si f est intégrable sur I et si K est un segment inclus dans I_1 (par exemple) on a aussi $K \subset I$ et par conséquent $\int_K f \leqslant \int_I f$, ce qui prouve l'intégrabilité de f sur I_1 (et l'inégalité

$\int_{I_1} f \leqslant \int_I f$). Si réciproquement f est intégrable sur I_1 et I_2, montrons que f est intégrable sur I. On peut supposer les intervalles numérotés pour que I_1 soit "à gauche" de I_2. Si $I_1 = \bigcup_{\nearrow} K_n^1$ et $I_2 = \bigcup_{\nearrow} K_n^2$ sont des écritures de ces deux intervalles comme réunion croissante de segments, on voit facilement que les suites

$$\alpha_n = \sup K_n^1 \quad \text{et} \quad \beta_n = \inf K_n^2$$

sont deux suites adjacentes et que la suite de segments

$$K_n = K_n^1 \cup [\alpha_n, \beta_n] \cup K_n^2$$

vérifie $I = \bigcup_{\nearrow} K_n$. On a alors

$$\int_{K_n} f = \int_{K_n^1} f + \int_{\alpha_n}^{\beta_n} f(t)dt + \int_{K_n^2} f$$

La première et la troisième des ces intégrales tendent respectivement vers $\int_{I_1} f$ et $\int_{I_2} f$ d'après la section précédente. La seconde vérifie

$$\left| \int_{\alpha_n}^{\beta_n} f(t)dt \right| \leqslant (\beta_n - \alpha_n) \sup_{[\alpha_1, \beta_1]} |f|$$

et tend donc vers 0 quand $n \to +\infty$. Le théorème (**12-1**.5) montre que f est intégrable sur I et que

$$\int_I f = \int_{I_1} f + \int_{I_2} f \quad \blacksquare$$

Il est donc naturel, pour que l'égalité précédente soit toujours valable de poser

$$\int_{\emptyset} f = 0$$

ce qui se justifie également par le résultat suivant :

COROLLAIRE 12-1.7 Si $f \in C_m^+(I)$ est intégrable sur I elle l'est également sur tout sous intervalle J de I et on a

$$\int_J f = \int_I f.1_J$$

où 1_J représente la fonction caractéristique de J.

Démonstration : Remarquons d'abord que la fonction $f.1_J$ est continue par morceaux sur I comme produit de deux telles fonctions. Elle est évidemment positive et intégrable sur I d'après la domination

$$f.1_J \leqslant f$$

Si $K \subset J$ est un segment

$$\int_K f = \int_K f.1_J \leqslant \int_I f.1_J$$

et donc f est intégrable sur J (ceci a en fait déja été vu dans la démonstration du théorème précédent). On a

$$\int_J f \leqslant \int_I f.1_J$$

Pour démontrer l'égalité, il suffit de remarquer que si $I = (J_1) \cup J \cup (J_2)$ est une partition de I en 1, 2 ou 3 intervalles non vides (faire un dessin) on a, par additivité par rapport à l'intervalle d'intégration, puis en utilisant la majoration précédente et enfin par linéarité de l'intégrale

$$\begin{aligned}\int_I f = \int_{J_1} f + \int_J f + \int_{J_2} f \\ \leqslant \int_I f.1_{J_1} + \int_I f.1_J + \int_I f.1_{J_2} \\ = \int_I f.(1_{J_1} + 1_J + 1_{J_2}) = \int_I f\end{aligned}$$

ce qui montre que l'inégalité précédente ne peut être stricte. ■

En particulier, comme pour $a \in I$ et $f \in \mathcal{C}_m^+(I)$ on a évidemment $\displaystyle\int_{\{a\}} f = 0$, l'intégrabilité sur $I \cap [a, +\infty[$ équivaut à l'intégrabilité sur $I \cap]a, +\infty[$ (intervalle éventuellement vide) et on a également le

COROLLAIRE 12-1.8 Si $f \in \mathcal{C}_m^+(I)$ et $a \in I$ est quelconque, f est intégrable sur I si et seulement si f l'est sur chacun des intervalles $I \cap]-\infty, a]$ et $I \cap [a, +\infty[$. Si c'est le cas, on a

$$\int_I f = \int_{I \cap]-\infty, a]} f + \int_{I \cap [a, +\infty[} f$$

L'intérêt de ce corollaire est de pouvoir ramener l'étude de l'intégrabilité d'une fonction continue par morceaux positive sur un intervalle ouvert à l'étude de l'intégrabilité sur des intervalles semi-ouverts (on laisse de côté le cas trivial des segments) : c'est très souvent ce qu'on fait en pratique, l'étude sur $]a, b[$ pouvant se dissocier en études sur $[c, b[$ et $]a, c]$, pour un choix arbitraire de $c \in]a, b[$. Le théorème qui suit montre que cela se fera dans l'immense majorité des cas par une étude locale de la fonction f en a et en b. Enfin, de nombreux théorèmes seront énoncés pour une fonction définie sur un intervalle de la forme $[a, b[$ avec $b \in \mathbb{R} \cup \{+\infty\}$, le cas symétrique $]a, b]$ avec éventuellement $a = -\infty$ ne nécessitant qu'une petite modification technique laissée aux soins du lecteur attentif.

COROLLAIRE 12-1.9 Si $f \in \mathcal{C}_m^+([a, b[)$, et $c \in [a, b[$, f est intégrable sur $[a, b[$ ssi elle l'est sur $[c, b[$ et, si c'est le cas,

$$\int_{[a,b[} f = \int_a^c f(t)dt + \int_{[c,b[} f$$

Démonstration : On sait en effet que f étant en particulier continue par morceaux sur $[a,c]$ est intégrable sur $[a,c[$ et d'intégrale égale à $\int_{[a,c]} f$. ■

Comme c peut être choisi très proche de b, seul importe finalement le comportement de f au voisinage de b.

EXERCICE 12-1.10 Montrer que, si $f \in \mathcal{C}_m^+(I)$ et A et B sont deux sous-intervalles de I d'intersection non vide, on a

$$\int_A f + \int_B f = \int_{A\cup B} f + \int_{A\cap B} f$$

12-1.4 Caractérisation de l'intégrabilité sur $[a,b[$

12-1.4.1 Théorème de base

Le théorème suivant est fondamental pour l'étude de l'intégrabilité des fonctions positives.

THÉORÈME 12-1.11 Si $f \in \mathcal{C}_m^+([a,b[)$, f est intégrable sur $[a,b[$ ssi la fonction

$$F(x) = \int_a^x f(t)dt$$

est majorée sur $[a,b[$. Si c'est le cas on a

$$\int_{[a,b[} f = \lim_{\substack{x\to b \\ x<b}} \int_a^x f(t)dt$$

Démonstration : La fonction F est évidemment croissante sur $[a,b[$. Dire que F est majorée sur $[a,b[$ équivaut à affirmer l'existence d'une limite finie pour F à gauche en b. Si $(x_n)_{n\in\mathbb{N}}$ est une suite croissante de points de $[a,b[$ convergente vers b, cela équivaut également au fait que la suite $(F(x_n))_{n\in\mathbb{N}}$ est bornée. L'équivalence avec l'intégrabilité de f est donc conséquence du théorème (**12-1**.5) en travaillant avec la suite de segments $K_n = [a,x_n]$. ■

Autre notation : Si f est intégrable (positive pour le moment) sur un intervalle (a,b) (ouvert, semi-ouvert ou fermé), son intégrale est aussi notée

$$\int_{(a,b)} f = \int_a^b f(t)dt$$

Cette notation ne présente pas d'ambiguïté puisque, si l'intervalle est (par exemple) de la forme $[a,b[$, on sait que $\int_{[a,b[} f = \int_{]a,b[} f$ (par additivité par rapport à l'intervalle d'intégration). De plus, si l'intervalle est un segment, on retrouve l'intégrale d'une fonction continue par morceaux sur ce segment.

12-1.4.2 Quelques fonctions "de référence"

Le théorème précédent permet de démontrer les résultats suivants :

- La fonction $t \mapsto \frac{1}{t^\alpha}$ est intégrable sur $[1, +\infty[$ ssi $\alpha > 1$ et on a alors

$$\int_1^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{\alpha - 1}$$

- La fonction $t \mapsto \frac{1}{t^\alpha}$ est intégrable sur $]0, 1]$ ssi $\alpha < 1$ et on a alors

$$\int_0^1 \frac{dt}{t^\alpha} = \frac{1}{1 - \alpha}$$

  Par translation, on en déduit que la condition nécessaire et suffisante d'intégrabilité de $t \mapsto \frac{1}{(t-a)^\alpha}$ (resp. $t \mapsto \frac{1}{(a-t)^\alpha}$) sur $]a, b]$, $b > a$ (resp. $[b, a[$, $b < a$) est aussi $\alpha < 1$.

- La fonction $t \mapsto -\ln t$ est intégrable sur $]0, 1]$ et $\int_{]0,1]} -\ln t \, dt = 1$. Cette fonction n'est pas intégrable sur $]0, +\infty[$.

- La fonction $t \mapsto e^{-at}$, avec $a > 0$, est intégrable sur $[0, +\infty[$ et

$$\int_0^{+\infty} e^{-at} \, dt = \frac{1}{a}$$

- La fonction $t \mapsto e^{-t^2}$ est intégrable sur $\mathbb{R}$. Pour des raisons de parité, il suffit de démontrer l'intégrabilité sur $[0, +\infty[$, qui est par exemple conséquence de la domination, sur $[1, +\infty[$, $e^{-t^2} \leqslant e^{-t}$. Un exercice traité lors de l'étude des intégrales sur un segment dépendant d'un paramètre (utilisant le calcul de la dérivée de la fonction $x \mapsto \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2} dt$) donne par ailleurs

$$\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2} \text{ et donc } \int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$$

12-1.5 Utilisation de critères de comparaison

La connaissance de l'intégrabilité (ou la non-intégrabilité) de certaines fonctions de référence est très utile, car par majorations ou minorations on pourra en déduire l'existence ou la non-existence de très nombreuses intégrales. Les théorèmes suivants procèdent tous de cette idée. On notera l'analogie avec la théorie des séries à termes positifs.

THÉORÈME 12-1.12 Si f et g sont continues par morceaux : $[a, b[\to \mathbb{R}^+$ et si $f \leqslant g$ au voisinage de b, alors

$$g \text{ intégrable sur } [a, b[\Longrightarrow f \text{ intégrable sur } [a, b[$$

Démonstration : il suffit de choisir $c \in [a,b[$ tel que la majoration $f \leqslant g$ soit valable sur $[c,b[$ et de remarquer ensuite que g étant intégrable sur $[c,b[$, il en est de même de f par domination. ■

On en déduit bien sûr que, si f n'est pas intégrable sur $[a,b[$, il en est de même de g.

COROLLAIRE 12-1.13 (Domination, négligeabilité) Si f et g sont positives continues par morceaux sur $[a,b[$ et si $f \underset{b}{=} O(g)$ (donc a fortiori si $f \underset{b}{=} o(g)$) l'intégrabilité de g sur $[a,b[$ entraîne celle de f.

COROLLAIRE 12-1.14 (Equivalence) Si f et g sont continues par morceaux de $[a,b[$ dans $\mathbb{R}^+$ et si $f \underset{b}{\sim} g$ alors

$$g \text{ intégrable sur } [a,b[\iff f \text{ intégrable sur } [a,b[$$

EXEMPLE 12-1.15 (Intégrales de Bertrand) La fonction $t \mapsto \dfrac{1}{t^\alpha |\ln t|^\beta}$ est intégrable sur l'intervalle $[2,+\infty[$ si et seulement si $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$). Donner de même une condition nécessaire et suffisante d'intégrabilité sur $]1,+\infty[$, puis sur $\left]0, \dfrac{1}{2}\right]$.

EXEMPLE 12-1.16 La fonction $f(t) = \dfrac{1}{\sin t + t^2}$ est-elle intégrable sur $\left]0, \dfrac{\pi}{2}\right]$?

EXEMPLE 12-1.17 Discuter en fonction de α et $\beta \in \mathbb{R}$ l'intégrabilité sur $]1,+\infty[$ de la fonction

$$f(x) = \frac{1}{|x^\alpha - 1|^\beta}$$

EXEMPLE 12-1.18 Si l'intervalle $[a,b[$ est **borné** et si $f \in \mathcal{C}_m^+([a,b[)$ est bornée, elle est intégrable (c'est en particulier le cas si f possède **une limite finie en** b, f est alors prolongeable en une fonction continue par morceaux sur $[a,b]$). Dans ce cas on a la majoration

$$0 \leqslant \int_a^b f(t)\,dt \leqslant (b-a)\,\|f\|_\infty$$

Attention ! Si $f : [a,+\infty[\to \mathbb{R}^+$ est continue par morceaux et possède une limite l en $+\infty$, cette limite est forcément nulle (car sinon $f \geqslant \dfrac{l}{2}$ au voisinage de $+\infty$ et ne peut être intégrable). Mais f peut très bien être intégrable sans tendre vers 0 à l'infini. La situation n'est pas la même que pour les séries. Considérer par exemple la fonction définie sur $[1,+\infty[$ (et en escalier sur tout segment)

$$f = \sum_{n=1}^{+\infty} n \mathbf{1}_{[n,n+\frac{1}{n^3}]}$$

12-1.6 "Intégration" des relations de comparaison

Continuons l'analogie avec les séries à termes positifs pour obtenir des résultats correspondant aux théorèmes de sommation des relations de comparaison.

THÉORÈME 12-1.19 (cas des fonctions intégrables) Soient deux fonctions φ et $\psi \in C_m^+([a,b[)$ avec ψ intégrable sur $[a,b[$. Alors

$$\begin{array}{lll} \textbf{si } \varphi(x) \underset{x\to b}{=} O(\psi(x)) & \textbf{alors} & \displaystyle\int_x^b \varphi(t)\,dt \underset{x\to b}{=} O\left(\int_x^b \psi(t)\,dt\right) \\ \textbf{si } \varphi(x) \underset{x\to b}{=} o(\psi(x)) & \textbf{alors} & \displaystyle\int_x^b \varphi(t)\,dt \underset{x\to b}{=} o\left(\int_x^b \psi(t)\,dt\right) \\ \textbf{si } \varphi(x) \underset{x\to b}{\sim} \psi(x) & \textbf{alors} & \displaystyle\int_x^b \varphi(t)\,dt \underset{x\to b}{\sim} \int_x^b \psi(t)\,dt \end{array}$$

Démonstration : remarquons que les hypothèses entraînent l'intégrabilité de φ. Supposons par exemple $\varphi \underset{b}{\sim} \psi$. Si $\varepsilon > 0$ est fixé arbitrairement, il existe $c \in [a,b[$ tel que

$$\forall t \in [c,b[\quad (1-\varepsilon)\psi(t) \leqslant \varphi(t) \leqslant (1+\varepsilon)\psi(t)$$

(remarquer l'importance de la positivité de ψ). On a alors, pour $x \geqslant c$

$$(1-\varepsilon)\int_x^b \psi(t)\,dt \leqslant \int_x^b \varphi(t)\,dt \leqslant (1+\varepsilon)\int_x^b \psi(t)\,dt$$

ce qui démontre l'équivalence en b des fonctions $\int_x^b \varphi(t)\,dt$ et $\int_x^b \psi(t)\,dt$.

■

EXERCICE 12-1.20 Existence et équivalent, pour $x \to +\infty$ de

$$\int_x^{+\infty} \frac{t+\frac{1}{t}}{t^2 \ln^2 t + \sin t}\,dt$$

THÉORÈME 12-1.21 (majorante non intégrable) Soient φ et $\psi \in C_m^+([a,b[)$ avec ψ non intégrable sur $[a,b[$. Alors $\lim\limits_{x\to b} \int_a^x \psi(t)\,dt = +\infty$ et

$$\begin{array}{lll} \textbf{si } \varphi(x) \underset{x\to b}{=} O(\psi(x)) & \textbf{alors} & \displaystyle\int_a^x \varphi(t)\,dt \underset{x\to b}{=} O\left(\int_a^x \psi(t)\,dt\right) \\ \textbf{si } \varphi(x) \underset{x\to b}{=} o(\psi(x)) & \textbf{alors} & \displaystyle\int_a^x \varphi(t)\,dt \underset{x\to b}{=} o\left(\int_a^x \psi(t)\,dt\right) \\ \textbf{si } \varphi(x) \underset{x\to b}{\sim} \psi(x) & \textbf{alors} & \displaystyle\int_a^x \varphi(t)\,dt \underset{x\to b}{\sim} \int_a^x \psi(t)\,dt \end{array}$$

Démonstration : remarquons que les hypothèses de domination ou de négligeabilité n'apportent pas d'information sur l'intégrabilité de φ, mais la conclusion du théorème n'a vraiment d'intérêt que lorsque φ n'est pas intégrable (on compare alors deux infiniment grands au

voisinage de b). Supposons par exemple $\varphi \underset{b}{=} o(\psi)$. Si $\varepsilon > 0$ est fixé arbitrairement, il existe $c \in [a, b[$ tel que

$$\forall t \in [c, b[\quad \varphi(t) \leqslant \varepsilon\, \psi(t)$$

si $c \leqslant x < b$, on a

$$\int_a^x \varphi(t)\,dt = \int_a^c \varphi(t)\,dt + \int_c^x \varphi(t)\,dt \leqslant \int_a^c \varphi(t)\,dt + \varepsilon \int_a^x \psi(t)\,dt$$

Comme $\lim\limits_{x\to b} \int_a^x \psi(t)\,dt = +\infty$, supposons $\int_a^x \psi(t)\,dt > 0$ pour $x \geqslant c$. On a

$$\lim_{x\to b} \frac{\int_a^c \varphi(t)\,dt}{\int_a^x \psi(t)\,dt} = 0$$

et on peut donc trouver $c_1 \in [c, b[$ tel que, pour tout $x \in [c_1, b[$ on ait

$$\frac{\int_a^c \varphi(t)\,dt}{\int_a^x \psi(t)\,dt} \leqslant \varepsilon$$

On a donc

$$\forall x \in [c_1, b[\quad 0 \leqslant \frac{\int_a^x \varphi(t)\,dt}{\int_a^x \psi(t)\,dt} \leqslant 2\varepsilon$$

ce qui démontre bien que $\int_a^x \varphi(t)\,dt \underset{x\to b}{=} o\left(\int_a^x \varphi(t)\,dt\right)$. ∎

EXERCICE 12-1.22 Trouver un développement asymptotique, pour $x \to +\infty$, de $\int_0^x e^{t^2}\,dt$ (utiliser des intégrations par parties).

EXERCICE 12-1.23 Montrer que $\int_e^x \frac{dt}{\ln t} \underset{+\infty}{\sim} \frac{x}{\ln x}$

12-1.7 Théorème de convergence monotone

Dans la théorie de l'intégrale sur un segment, l'outil fondamental pour "passer à la limite sous le signe intégrale" est la convergence uniforme, ceci étant conséquence de l'inégalité

$$\left| \int_a^b f_n(t)\,dt - \int_a^b f(t)\,dt \right| \leqslant (b-a)\, \|f_n - f\|_\infty$$

Cet outil reste valable dans la théorie que nous présentons ici, pour peu que l'intervalle d'intégration soit **borné**. En effet, si les fonctions $(f_n)_{n\in\mathbb{N}}$ sont continues

par morceaux (et positives pour le moment) intégrables sur un intervalle (a,b) borné et convergent uniformément sur cet intervalle vers une fonction f supposée continue par morceaux, l'inégalité

$$0 \leqslant f \leqslant f_n + \|f - f_n\|_\infty 1_{(a,b)}$$

prouve l'intégrabilité de f. On a facilement (par passage à la limite en b)

$$\left| \int_a^b f_n(t)\,dt - \int_a^b f(t)\,dt \right| \leqslant (b-a)\,\|f_n - f\|_\infty$$

Cette inégalité montre ainsi que la convergence uniforme entraîne la convergence des intégrales **sur un intervalle borné**.

La convergence uniforme est cependant une notion techniquement compliquée, qui de plus n'est pas vraiment adaptée au cas où l'intervalle d'intégration n'est plus borné. Considérer par exemple le cas

$$I = [0, +\infty[\quad \text{et} \quad f_n = \frac{1}{n} 1_{[0,n]}$$

Le théorème qui suit est fondamental. Nous l'énonçons dans le cadre des fonctions positives qui est vraiment la base de la théorie. Il se généralisera ultérieurement au cas des fonctions réelles. L'idée de passage à la limite sur des intervalles croissants utilisée pour construire l'intégrale trouve ici une extension remarquable.

THÉORÈME 12-1.24 (convergence monotone) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathcal{C}_m^+(I)$, croissante ($f_n \leqslant f_{n+1}$) qui converge simplement sur I vers une fonction $f \in \mathcal{C}_m^+(I)$. La fonction f est intégrable sur I ssi la suite des intégrales $\left(\int_I f_n\right)_{n\in\mathbb{N}}$ est majorée. Si c'est le cas, on a

$$\int_I f = \sup_{n\in\mathbb{N}} \int_I f_n = \lim_{n\to+\infty} \int_I f_n$$

Démonstration : la condition est évidemment nécessaire. En effet si f est intégrable, comme pour tout n on a $f_n \leqslant f$ on a aussi

$$\forall n \in \mathbb{N} \quad \int_I f_n \leqslant \int_I f$$

et en conséquence

$$\sup_{n\in\mathbb{N}} \int_I f_n \leqslant \int_I f$$

Il reste à démontrer que, si $M = \sup_{n\in\mathbb{N}} \int_I f_n$ est fini alors f est intégrable d'intégrale égale à la limite des intégrales des f_n.

• Dans le cas particulier où la suite de fonctions f_n converge uniformément vers f sur tout segment inclus dans I, la démonstration est très simple : dans ce cas, si K est un segment inclus dans I on a

$$\int_K f_n \leqslant \int_I f_n$$

Le terme de gauche tend vers $\int_K f$ (convergence uniforme). La suite de droite est croissante et majorée, donc convergente vers M. On a alors

$$\forall K \text{ segment } \subset I \quad \int_K f \leqslant \sup_{n\in\mathbb{N}} \int_I f_n$$

ce qui démontre l'intégrabilité de f et l'inégalité

$$\int_I f \leqslant \sup_{n\in\mathbb{N}} \int_I f_n$$

Comme l'inégalité inverse a été vue plus haut, le théorème en découle dans ce cas particulier.

• La démonstration générale est basée sur deux lemmes :

LEMME 12-1.25 (Théorème de Dini) Si $(j_n)_{n\in\mathbb{N}}$ est une suite monotone de fonctions continues de $[a,b]$ dans $\mathbb{R}$ qui converge simplement sur $[a,b]$ vers une fonction continue j, alors la convergence est uniforme.

En effet, quitte à multiplier par -1, on peut supposer la suite j_n croissante ($j_n \leqslant j_{n+1}$) vers j. Si $\varepsilon > 0$ est choisi arbitrairement

$$K_n = \{x \in [a,b] \mid j(x) - j_n(x) \geqslant \varepsilon\ \}$$

est un fermé de $[a,b]$ donc est compact. On a clairement $K_{n+1} \subset K_n$ et $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$. La propriété de Borel-Lebesgue entraîne l'existence d'un entier n_0 tel que $K_{n_0} = \emptyset$. On a donc

$$\forall n \geqslant n_0 \quad \forall x \in [a,b] \quad 0 \leqslant j(x) - j_n(x) < \varepsilon$$

ce qui prouve la convergence uniforme. ■

LEMME 12-1.26 (convergence bornée) Si (g_n) est une suite de fonctions continues par morceaux d'un segment $[a,b]$ dans $\mathbb{R}^+$ convergeant simplement vers la fonction nulle et telle qu'il existe une fonction continue par morceaux $\varphi : [a,b] \to \mathbb{R}^+$ vérifiant

$$\forall n \quad 0 \leqslant g_n \leqslant \varphi \quad \textbf{sur} \quad [a,b]$$

alors

$$\lim_{n\to+\infty} \int_a^b g_n(t)\,dt = 0$$

Démonstration : si ce n'est pas le cas,

$$\exists \alpha > 0 \quad \forall N \in \mathbb{N} \quad \exists n \geqslant N \quad \int_a^b g_n(t)\,dt > \alpha$$

On peut donc extraire de la suite (g_n) une suite $f_n = g_{\psi(n)}$ avec

$$\forall n \quad \int_a^b f_n(t)\,dt > \alpha$$

Comme f_n est continue par morceaux positive, on peut trouver une fonction continue h_n sur $[a,b]$ vérifiant

$$0 \leqslant h_n \leqslant f_n \quad \text{et} \quad \int_a^b (f_n(t) - h_n(t))\,dt \leqslant \frac{\alpha}{2}$$

(il suffit de trouver une fonction en escalier s_n vérifiant $0 \leqslant s_n \leqslant f_n$ approchant suffisamment f_n en norme uniforme, puis d'approcher s_n "par en dessous" par une fonction continue h_n affine par morceaux, faire un dessin !).

Les fonctions h_n sont donc continues, positives, dominées par φ et convergent simplement vers 0 en vérifiant

$$\forall n \quad \int_a^b h_n(t)\,dt > \beta = \frac{\alpha}{2} \qquad (1)$$

Pour n fixé, la suite $\left(l_p^n\right)_{p\in\mathbb{N}}$ définie par

$$l_p^n = \sup(h_n, h_{n+1}, \dots, h_{n+p})$$

est une suite **croissante** de fonctions **continues** (sup. d'un nombre **fini** de fonctions continues), dont la suite des intégrales $\left(\int_a^b l_p^n(t)\,dt\right)_{p\in\mathbb{N}}$ est croissante, majorée par $\int_a^b \varphi(t)\,dt$, et donc convergente vers une limite λ_n vérifiant

$$\beta < \lambda_n = \lim_{p\to+\infty} \int_a^b l_p^n(t)\,dt \leqslant \int_a^b \varphi(t)\,dt$$

On peut donc trouver un rang $p(n)$ tel que

$$\forall k \geqslant p(n) \quad \lambda_n \geqslant \int_a^b l_k^n(t)\,dt \geqslant \lambda_n - \frac{\beta}{2^{n+1}} \qquad (2)$$

Posons $i_n = l_{p(n)}^n$. C'est une suite de fonctions continues sur $[a,b]$ qui converge simplement vers 0. En effet, si $t \in [a,b]$, comme la suite $(h_k(t))_{k\in\mathbb{N}}$ converge vers 0, si on se donne $\varepsilon > 0$ on peut trouver un rang k_0 tel que $k \geqslant k_0 \Longrightarrow h_k(t) \leqslant \varepsilon$. On a alors

$$\forall n \geqslant k_0 \quad \forall p \in \mathbb{N} \quad 0 \leqslant l_p^n(t) \leqslant \varepsilon$$
$$\text{et en particulier} \quad \forall n \geqslant k_0 \quad i_n(t) \leqslant \varepsilon$$

Posons (enfin !)

$$j_n = \inf(i_1, \dots, i_n)$$

On forme ainsi une suite de fonctions continues, décroissante et qui converge simplement vers 0 sur $[a,b]$ (puisque $0 \leqslant j_n \leqslant i_n$). D'après le théorème de Dini cité plus haut, la convergence est uniforme et en particulier

$$\lim_{n\to+\infty} \int_a^b j_n(t)\,dt = 0 \qquad (3)$$

Pour $t \in [a,b]$, il existe $k \in \{1,\dots,n\}$ tel que

$$0 \leqslant i_n(t) - j_n(t) = i_n(t) - i_k(t)$$

avec

$$i_n(t) - i_k(t) = \sup(h_n(t), h_{n+1}(t), \dots, h_{n+p(n)}(t)) - i_k(t)$$

En particulier, on obtient

$$0 \leqslant i_n(t) - j_n(t) \leqslant \sup(h_k(t), h_{k+1}(t), \dots, h_{n+k+p(n)+p(k)}(t)) - i_k(t)$$

On a donc majoré la différence $i_n(t) - j_n(t)$ par

$$l^k_{n+p(n)+p(k)}(t) - l^k_{p(k)}(t) = \alpha_k(t)$$

où α_k est positive et d'intégrale plus petite que $\dfrac{\beta}{2^{k+1}}$ d'après les encadrements (2). Comme dans le cas $k = n$ la différence est nulle, on a donc

$$0 \leqslant i_n(t) - j_n(t) \leqslant \sum_{k=1}^{n-1} \alpha_k(t)$$

ce qui donne en intégrant

$$\int_a^b j_n(t)\,dt \geqslant \int_a^b i_n(t)\,dt - \sum_{k=1}^{n-1} \frac{\beta}{2^{k+1}}$$

Comme $i_n \geqslant h_n$ on a d'après (1)

$$\int_a^b j_n(t)\,dt \geqslant \beta - \sum_{k=1}^{n-1} \frac{\beta}{2^{k+1}} > \frac{\beta}{2}$$

ce qui contredit (3). ■

Ces deux lemmes donnent immédiatement la démonstration du théorème de convergence monotone dans le cas général :

Si $K \subset I$ est un segment, avec les notations de l'énoncé du théorème, la suite de fonctions

$$g_n = (f - f_n)|_K$$

est simplement convergente vers 0 sur K et est dominée par la fonction $f|_K$. On a donc, d'après le lemme précédent

$$\int_K f = \lim_{n\to+\infty} \int_K f_n \leqslant \sup_n \int_I f_n$$

ce qui prouve l'intégrabilité de f avec $\displaystyle\int_I f \leqslant \sup_n \int_I f_n$, mais l'inégalité inverse a déjà été vue. ■

12-1.8 Exercices d'application

La démonstration précédente est très technique ! Concentrons-nous à présent sur les applications :

EXERCICE 12-1.27 Calculer

$$\lim_{n\to+\infty}\int_0^n (1-\frac{x}{n})^n e^{\frac{x}{2}}\,dx$$

EXERCICE 12-1.28 Calculer

$$\lim_{n\to+\infty}\left[\left(\frac{1}{n}\right)^n+\left(\frac{2}{n}\right)^n+\ldots+\left(\frac{n-1}{n}\right)^n\right]$$

Indication : il est d'abord intéressant d'écrire les plus grands termes en premier ! On remarque alors que la suite à étudier est intégrale sur $[0,+\infty[$ de la fonction en escalier

$$s_n(t)=\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)^n 1_{[k,k+1[}(t)$$

Appliquer le théorème de convergence monotone à cette suite. On retrouve pratiquement l'exercice précédent. ■

12-1.9 Intégration terme à terme d'une série de fonctions positives

Le théorème qui suit est conséquence immédiate du théorème de convergence monotone.

THÉORÈME 12-1.29 Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions de $C_m^+(I)$ intégrables, telles que la série de fonctions

$$S(x)=\sum_{n=0}^{+\infty}u_n(x)$$

converge simplement sur I, sa somme étant continue par morceaux sur I. Pour que S soit intégrable sur I, il faut et il suffit que la série $\sum_{n=0}^{+\infty}\int_I u_n$ soit convergente, et on a alors

$$\int_I S=\int_I\sum_{n=0}^{+\infty}u_n=\sum_{n=0}^{+\infty}\int_I u_n$$

Démonstration : il suffit d' appliquer le théorème de convergence monotone à la suite de fonctions $f_n=\sum_{k=0}^{n}u_k$. ■

EXERCICE 12-1.30 Démontrer que

$$\int_0^1 \frac{\ln x}{x-1}dx = \sum_{n=1}^{+\infty} \frac{1}{n^2}$$

REMARQUE 12-1.31 Pour illustrer l'analogie entre la théorie de l'intégrale et celle des familles sommables, on peut remarquer que le théorème précédent "contient" le théorème de Fubini sur l'interversion des sommations pour les suites doubles de réels positifs. Si $(u_{n,m})_{(n,m)\in\mathbb{N}^2}$ est une suite double de réels positifs tels que $\left(\sum_{n=0}^{+\infty}\left(\sum_{m=0}^{+\infty} u_{n,m}\right)\right)$ converge, il est facile de voir, en définissant les fonctions sur $[0,+\infty[$

$$u_n = \sum_{m=0}^{+\infty} u_{n,m} 1_{[m,m+1[}$$

que $S = \sum_{n=0}^{+\infty} u_n = \sum_{m=0}^{+\infty}\left(\sum_{n=0}^{+\infty} u_{n,m}\right) 1_{[m,m+1[}$ est continue par morceaux sur $[0,+\infty[$ (la seule difficulté, une fois comprises les notations, est en fait de prouver la convergence des séries $\sum_{n=0}^{+\infty} u_{n,m}$). L'hypothèse faite sur la suite double correspond exactement à la convergence de la série des intégrales des u_n. Le théorème sur les suites doubles est alors conséquence facile du théorème d'intégration terme à terme qui précède.

12-2 Fonctions complexes intégrables

12-2.1 Intégrale d'une fonction complexe

12-2.1.1 Intégrabilité. Espace vectoriel des fonctions intégrables

DÉFINITION 12-2.1 *Une fonction continue par morceaux $f : I \to \mathbb{C}$ est dite intégrable (sur I) si et seulement si $|f|$ l'est.*

Le théorème suivant montre que l'on peut toujours se ramener à des fonctions à valeurs réelles :

THÉORÈME 12-2.2 Une fonction $f \in C_m(I,\mathbb{C})$ est intégrable si et seulement ses parties réelle et imaginaire le sont.

Démonstration : conséquence immédiate des inégalités

$$|\mathrm{Re}\, f| \leqslant |f|, \quad |\mathrm{Im}\, f| \leqslant |f| \quad \text{et} \quad |f| \leqslant |\mathrm{Re}\, f| + |\mathrm{Im}\, f| \quad \blacksquare$$

Dans le cas des fonctions réelles, on peut ramener l'intégrabilité de f à celle de ses parties positives et négatives :

DÉFINITION 12-2.3 *Si $f : I \to \mathbb{R}$ est une fonction, on note*

$$f^+ = \max(f, 0) = \frac{1}{2}(|f| + f) \quad et \quad f^- = \max(-f, 0) = \frac{1}{2}(|f| - f)$$

ces fonctions (positives !) sont appelées respectivement ***partie positive*** *et* ***partie négative*** *de f. Si f est continue, il en est de même de f^+ et f^-. Conclusion analogue pour une fonction continue par morceaux. On a évidemment*

$$f = f^+ - f^- \quad et \quad |f| = f^+ + f^-$$

THÉORÈME 12-2.4 Si $f : I \to \mathbb{R}$ est continue par morceaux, elle est intégrable si et seulement si f^+ et f^- le sont.

Démonstration : conséquence immédiate des inégalités

$$f^+ \leqslant |f|, \quad f^- \leqslant |f| \quad \text{et} \quad |f| = f^+ + f^- \quad \blacksquare$$

Comme la définition ramène à l'étude d'une fonction positive, l'intégrabilité d'une fonction $f \in \mathcal{C}_m(I, \mathbb{K})$ se prouve en général par domination : s'il existe une fonction intégrable $\varphi \in \mathcal{C}_m^+(I)$ vérifiant $|f| \leqslant \varphi$, alors f est intégrable. Rappelons que, si I est de la forme $[a, b[$, il suffit d'avoir la majoration précédente au voisinage de b.

EXEMPLE 12-2.5 La fonction $t \mapsto \dfrac{\sin t}{t^2 + t^2 \ln t}$ est intégrable sur $[1, +\infty[$. L'est-elle sur $]0, 1]$?

EXEMPLE 12-2.6 La fonction $t \mapsto \dfrac{e^{it}}{t^\alpha}$ est intégrable sur $[1, +\infty[$ si et seulement si $\alpha > 1$. On en déduit que, pour $\alpha \leqslant 1$, l'une au moins des fonctions $s_\alpha : t \mapsto \dfrac{\sin t}{t^\alpha}$ et $c_\alpha : t \mapsto \dfrac{\cos t}{t^\alpha}$ est non intégrable. Si l'on montre que la fonction s_1 est non intégrable, on en déduira (comment ?) qu'il en est de même de s_α. On peut montrer que la suite

$$S_n = \int_\pi^{n\pi} \frac{|\sin t|}{t} dt$$

n'est pas bornée, en l'écrivant comme somme partielle de la série de terme général

$$u_k = \int_{(k-1)\pi}^{k\pi} \frac{|\sin t|}{t} dt \underset{k \to +\infty}{\sim} \frac{2}{k\pi}$$

Un argument identique montrerait la non intégrabilité de c_α pour $\alpha \leqslant 1$.

THÉORÈME 12-2.7 (et définition) L'ensemble des fonctions continues par morceaux $I \to \mathbb{K}$ intégrables sur I est un $\mathbb{K}$-espace vectoriel, que l'on note $\mathcal{L}^1(I, \mathbb{K})$.

C'est en effet un sous-espace de $\mathcal{C}_m(I, \mathbb{K})$. La stabilité par homothétie est évidente, la stabilité pour l'addition résultant de la domination

$$|f + g| \leqslant |f| + |g| \quad \blacksquare$$

12-2.1.2 Intégrale d'une fonction complexe.

THÉORÈME 12-2.8 (et définition : intégrale) Si $f \in \mathcal{L}^1(I, \mathbb{K})$, et si $(K_n)_{n\in\mathbb{N}}$ est une suite croissante de segments de réunion égale à I, la suite des intégrales $\int_{K_n} f$ est convergente et sa limite ne dépend pas de la suite K_n choisie. On l'appelle intégrale de f sur I et on la note

$$\int_I f = \lim_{n\to+\infty} \int_{K_n} f \qquad \textbf{avec} \quad I = \bigcup_{\nearrow} K_n$$

Démonstration : en séparant partie réelle et imaginaire, on peut se ramener au cas où f est réelle. Il suffit d'écrire

$$\int_{K_n} f = \int_{K_n} f^+ - \int_{K_n} f^-$$

et d'appliquer aux fonctions positives intégrables f^+ et f^- les résultats des sections qui précèdent. ■

Les propriétés qui suivent sont conséquence immédiates de la définition, f représentant un élément quelconque de $\mathcal{L}^1(I, \mathbb{K})$.

- Si $\mathbb{K} = \mathbb{R}$, on a $\int_I f = \int_I f^+ - \int_I f^-$
- Si $\mathbb{K} = \mathbb{C}$, on a $\int_I f = \int_I \mathrm{Re}\, f + i \int_I \mathrm{Im}\, f$, et donc $\overline{\int_I f} = \int_I \overline{f}$
- $\left|\int_I f\right| \leqslant \int_I |f|$
- L'application $f \mapsto \int_I f$ est une forme linéaire sur $\mathcal{L}^1(I, \mathbb{K})$.

En particulier, lorsque $I = [a, b[$, situation à laquelle on peut toujours se ramener comme on l'a déja vu, on a le

COROLLAIRE 12-2.9 Si f est intégrable sur $[a, b[$, on a

$$\int_{[a,b[} f = \lim_{\substack{x\to b \\ x<b}} \int_a^x f(t)\, dt$$

Attention ! Lorsque f est réelle positive (ou, si on réfléchit un peu, de signe constant au voisinage de b), l'existence de cette limite entraîne réciproquement l'intégrabilité de f. **Ce n'est plus vrai dans le cas général.** Il suffit de méditer le calcul suivant :

EXEMPLE 12-2.10 (contre-exemple !) La fonction $f : t \mapsto \dfrac{\sin t}{t}$ est continue sur $[1, +\infty[$. Une intégration par parties donne

$$\forall x \geqslant 1 \quad \int_1^x \frac{\sin t}{t} dt = \cos 1 - \frac{\cos x}{x} - \int_1^x \frac{\cos t}{t^2} dt$$

et, la fonction $t \mapsto \dfrac{\cos t}{t^2}$ étant intégrable sur $[1, +\infty[$

$$\lim_{x \to +\infty} \int_1^x \frac{\sin t}{t} dt = \cos 1 - \int_{[1,+\infty[} \frac{\cos t}{t^2} dt$$

et pourtant, comme on l'a vu plus haut, f n'est pas intégrable sur $[1, +\infty[$. On reviendra ultérieurement sur ce genre de situation (*cf.* section **12-4**).

12-2.1.3 Intégration sur un sous-intervalle

Comme pour intégrer une fonction complexe on intègre en fait quatre fonctions positives, les résultats de la section (**12-1.3**) donnent immédiatement

THÉORÈME 12-2.11 Si $f \in \mathcal{L}^1(I, \mathbb{K})$, f est intégrable sur tout sous-intervalle J de I et on a

$$\int_J f = \int_I f 1_J$$

THÉORÈME 12-2.12 Si $f \in \mathcal{L}^1(I, \mathbb{K})$ et si $I = I_1 \cup I_2$ est une partition de I en deux intervalles, on a

$$\int_I f = \int_{I_1} f + \int_{I_2} f$$

Si f est une fonction continue par morceaux intégrable sur un intervalle $]a, b[$, on notera encore

$$\int_{]a,b[} f = \int_a^b f(t)\, dt$$

Cette notation est sans ambiguïté, puisque si par exemple $a \in \mathbb{R}$ et si f est continue par morceaux intégrable sur $[a, b[$, on a

$$\int_{]a,b[} f = \int_{[a,b[} f$$

Si $a < b$, on notera encore

$$\int_b^a f(t)\, dt = -\int_a^b f(t)\, dt$$

ce qui permet de généraliser la relation de Chasles.

12-2.2 Formule de changement de variable

THÉORÈME 12-2.13 (changement de variable) Soient I et J deux intervalles réels, φ une bijection de classe C^1 de $I \to J$ et f une fonction continue par morceaux de $J \to \mathbb{C}$. La fonction f est intégrable sur J si et seulement si la fonction $(f \circ \varphi)|\varphi'|$ est intégrable sur I et on a alors

$$\int_I (f \circ \varphi)|\varphi'| = \int_{J=\varphi(I)} f$$

Démonstration : on peut se ramener au cas où f est réelle positive. Remarquons que, comme φ est injective et continue, elle est strictement monotone, et la fonction φ' garde un signe constant sur I. Si K_n est une suite croissante de segments de réunion égale à I, la suite des $K_n' = \varphi(K_n)$ est une suite de segments de réunion égale à $J = \varphi(I)$. Si, par exemple, φ est décroissante on a

$$\forall n \in \mathbb{N} \qquad -\int_{K_n} (f \circ \varphi)\varphi' = \int_{K_n'} f$$

par la formule de changement de variable pour l'intégrale sur un segment, le signe $-$ provenant d'une interversion des bornes (les deux membres de l'égalité ont même signe). Dire que ces suites sont bornées revient à affirmer l'intégrabilité simultanée de f sur J et de $(f \circ \varphi)|\varphi'|$ sur I. La formule de changement de variable s'obtient alors en faisant tendre n vers $+\infty$. ■

12-2.3 Intégration des relations de comparaison

Ici encore, les théorèmes découlent immédiatement des résultats obtenus dans le cas positif. On notera cependant que la fonction de référence utilisée dans ces énoncés est positive. Les fonctions considérées sont évidemment continues par morceaux.

THÉORÈME 12-2.14 Si φ est intégrable positive sur $[a, b[$ et $f : [a, b[\to \mathbb{C}$ vérifie

$$f \underset{b}{=} O(\varphi) \quad \text{ou} \quad f \underset{b}{=} o(\varphi)$$

alors f est intégrable sur $[a, b[$ et on a respectivement

$$\int_x^b f(t)\,dt \underset{x \to b^-}{=} O\left(\int_x^b \varphi(t)\,dt\right) \quad \text{ou} \quad \int_x^b f(t)\,dt \underset{x \to b^-}{=} o\left(\int_x^b \varphi(t)\,dt\right)$$

Démonstration : l'intégrabilité de f est conséquence d'une domination $|f| \leqslant M\varphi$ au voisinage de b. La majoration

$$\left|\int_x^b f(t)\,dt\right| \leqslant \int_x^b |f(t)|\,dt$$

permet ensuite de revenir aux fonctions positives. ■

COROLLAIRE 12-2.15 Sous les mêmes hypothèses ($\varphi \geqslant 0$) et si $f \underset{b}{\sim} k\varphi$ avec $k \in \mathbb{C}^*$, on a aussi

$$\int_x^b f(t)\,dt \underset{x\to b^-}{\sim} k\int_x^b \varphi(t)\,dt$$

Démonstration : il suffit d'écrire $f = k\varphi + g$ avec $g \underset{b}{=} o(\varphi)$. Toutes les fonctions étant intégrables (par domination), on a

$$\int_x^b f(t)\,dt = k\int_x^b \varphi(t)\,dt + \int_x^b g(t)\,dt \underset{x\to b^-}{=} k\int_x^b \varphi(t)\,dt + o\left(\int_x^b \varphi(t)\,dt\right)$$

d'aprés le théorème précédent. ■

EXERCICE 12-2.16 Enoncer les résultats que l'on peut obtenir dans le cas d'une majorante non intégrable.

12-2.4 Convergence en moyenne, en moyenne quadratique

12-2.4.1 (Semi-)norme de la convergence en moyenne

THÉORÈME 12-2.17 Sur l'espace vectoriel $\mathcal{L}^1(I,\mathbb{K})$ l'application

$$f \mapsto \|f\|_1 = N_1(f) = \int_I |f|$$

est une semi-norme appelée semi-norme de la convergence en moyenne.

Elle a en effet toutes les propriétés d'une norme, sauf que la nullité de $\|f\|_1$ entraîne la nullité de f en tout point de continuité, ce qui veut dire que sur tout segment inclus dans I, f est nulle sauf en un nombre fini de points.

COROLLAIRE 12-2.18 Sur l'espace vectoriel $\mathcal{L}^1(I,\mathbb{K}) \cap \mathcal{C}^0(I,\mathbb{K})$, N_1 est une norme.

DÉFINITION 12-2.19 *Si $(f_n)_{n\in\mathbb{N}}$ et $f \in \mathcal{L}^1(I,\mathbb{K})$, on dit que la suite $(f_n)_{n\in\mathbb{N}}$* ***converge vers f en moyenne sur I*** *ssi*

$$\lim_{n\to+\infty} \|f - f_n\|_1 = 0$$

On notera qu'avec la définition que nous donnons ici, si on remplace la fonction f par une fonction g telle que $\|f-g\|_1 = 0$, on aura aussi convergence vers g. Ce n'est pas bien gênant. En tout cas, si on se limite à travailler dans l'espace normé $(\mathcal{L}^1(I,\mathbb{K}) \cap \mathcal{C}^0(I,\mathbb{K}), \|\ \|_1)$, il s'agit de la convergence pour une norme et il y a bien dans ce cas unicité de la limite.

Il est à noter que cette notion de convergence n'a pas grand rapport avec la convergence simple. Comme nous l'avons déja indiqué dans le cas de l'intégration sur un segment, il est facile de construire une suite de fonctions f_n convergeant en moyenne et telle que pour tout $t \in I$ la suite $(f_n(t))_{n\in\mathbb{N}}$ soit divergente! Dans le cas où **l'intervalle I est borné,** nous avons déjà évoqué le

THÉORÈME 12-2.20 Si I est borné et si $(f_n)_{n\in\mathbb{N}} \in \mathcal{L}^1(I,\mathbb{K})$ est une suite qui converge uniformément sur I vers une fonction f continue par morceaux, alors f est intégrable sur I et

$$\lim_{n\to+\infty} \|f - f_n\|_1 = 0$$

L'intérêt de la théorie de l'intégrale présentée ici est de fournir des théorèmes (assurant la convergence en moyenne) dont les hypothèses ne font pas appel à la notion de convergence uniforme, et qui s'appliqueront lorsque l'intervalle d'intégration est ou n'est pas borné. Nous généraliserons en effet le théorème de convergence monotone dans la section suivante.

L'inégalité du module montre qu'une convergence en moyenne entraîne la convergence des intégrales :

THÉORÈME 12-2.21 Si la suite $(f_n)_{n\in\mathbb{N}}$ converge en moyenne sur I vers f, on a en particulier

$$\lim_{n\to+\infty} \int_I f_n = \int_I f$$

La majoration $\left|\int_I f\right| \leqslant \|f\|_1$ montre également que l'intégrale est une forme linéaire continue sur l'espace normé $(\mathcal{L}^1(I,\mathbb{K}) \cap \mathcal{C}^0(I,\mathbb{K}), \| \|_1)$.

12-2.4.2 Fonctions de carré intégrable. Convergence en moyenne quadratique

DÉFINITION 12-2.22 *Une fonction $f \in \mathcal{C}^0_m(I,\mathbb{K})$ est dite de carré intégrable ssi $f^2 \in \mathcal{L}^1(I,\mathbb{K})$ (ce qui revient bien sûr, dans le cas complexe, à supposer l'intégrabilité de $|f|^2$).*

THÉORÈME 12-2.23 L'ensemble

$$\mathcal{L}^2(I,\mathbb{K}) = \left\{f \in \mathcal{C}^0_m(I,\mathbb{K}) \mid f \textbf{ est de carré intégrable}\right\}$$

est un espace vectoriel sur lequel

$$\|f\|_2 = N_2(f) = \sqrt{\int_I |f|^2}$$

définit une semi-norme (dite de la convergence en moyenne quadratique). On a véritablement une norme si on travaille en restriction à l'espace vectoriel $\mathcal{L}^2(I,\mathbb{K}) \cap \mathcal{C}^0(I,\mathbb{K})$.

Démonstration : on montre qu'on a un sous-espace vectoriel de $\mathcal{C}^0_m(I,\mathbb{K})$. La stabilité par homothétie est évidente. Pour l'addition, elle est conséquence de la domination

$$|f+g|^2 \leqslant (|f|+|g|)^2 \leqslant 2\left(|f|^2+|g|^2\right)$$

Seule reste à démontrer l'inégalité triangulaire (inégalité de Minkowski). Elle s'obtient par passage à la limite. Si $I = \bigcup\limits_{\nearrow} K_n$, on a

$$\forall n \quad \sqrt{\int_{K_n} |f+g|^2} \leqslant \sqrt{\int_{K_n} |f|^2} + \sqrt{\int_{K_n} |g|^2}$$

En faisant tendre n vers l'infini, on obtient bien

$$\|f+g\|_2 \leqslant \|f\|_2 + \|g\|_2 \quad \blacksquare$$

THÉORÈME 12-2.24 (Inégalité de Schwarz) Si f et $g \in \mathcal{L}^2(I, \mathbb{K})$, leur produit fg est intégrable et on a

$$\|fg\|_1 \leqslant \|f\|_2 \, \|g\|_2$$

en particulier

$$\left| \int_I fg \right| \leqslant \sqrt{\int_I |f|^2} \sqrt{\int_I |g|^2}$$

Démonstration : l'intégrabilité est conséquence de la majoration

$$|fg| \leqslant \frac{1}{2} \left(|f|^2 + |g|^2 \right)$$

et l'inégalité de Schwarz s'obtient également par un passage à la limite. ■

Il n'est pas étonnant de trouver ces inégalités. Nous verrons ultérieurement que l'application

$$(f, g) \mapsto \int_I fg$$

est un produit scalaire sur l'espace vectoriel $\mathcal{L}^2(I, \mathbb{R}) \cap C^0(I, \mathbb{R})$. Dans le cas complexe, c'est $\int_I \overline{f} g$ qui définit un produit scalaire hermitien.

EXERCICE 12-2.25 Dans le cas où l'intervalle I est **borné**, montrer que l'on a l'inclusion

$$\mathcal{L}^2(I, \mathbb{K}) \subset \mathcal{L}^1(I, \mathbb{K})$$

et que dans ce cas la convergence en moyenne quadratique entraîne la convergence en moyenne. Sur l'intervalle $]0,1]$ que dire de $f(x) = \frac{1}{\sqrt{x}}$? Sur $[1, +\infty[$ de $f(x) = \frac{1}{x}$?

12-3 Les théorèmes de convergence. Intégrales à paramètres

12-3.1 Théorème de convergence monotone

On l'énonce ici dans le cas de fonctions (réelles !) non nécessairement positives.

THÉORÈME 12-3.1 Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathcal{L}^1(I,\mathbb{R})$ qui converge simplement en croissant sur I vers une fonction f continue par morceaux. La fonction f est intégrable ssi

$$\sup_n \int_I f_n < +\infty$$

et on a alors

$$\int_I f = \lim_{n\to+\infty} \int_I f_n = \sup_n \int_I f_n$$

Démonstration : il suffit d'appliquer le théorème démontré dans le cas positif à la suite de fonctions $g_n = f_n - f_0$. ■

Dans le théorème qui suit, on s'affranchit de la monotonie.

12-3.2 Théorème de convergence dominée

THÉORÈME 12-3.2 Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathcal{L}^1(I,\mathbb{K})$ qui converge simplement sur I vers une fonction continue par morceaux f. On suppose qu'il existe une fonction $\varphi : I \to \mathbb{R}^+$ intégrable telle que

$$\forall n \in \mathbb{N} \quad |f_n| \leqslant \varphi$$

Alors la fonction f est intégrable, la suite f_n converge vers f en moyenne sur I et en particulier

$$\int_I f = \lim_{n\to+\infty} \int_I f_n$$

Démonstration : l'intégrabilité de f est évidemment conséquence de la majoration (obtenue par passage à la limite) $|f| \leqslant \varphi$.

• Dans le cas particulier où la convergence est uniforme sur tout segment inclus dans I, si $\varepsilon > 0$ est fixé arbitrairement, on peut trouver un segment $K \subset I$ tel que

$$\int_{I-K} \varphi < \frac{\varepsilon}{4}$$

($I - K$ est, selon les cas de figure, réunion d'un ou deux intervalles). Comme $|f - f_n| \leqslant 2\varphi$ on a alors

$$\forall n \quad \int_{I-K} |f - f_n| < \frac{\varepsilon}{2}$$

Comme la suite $|f - f_n|$ converge uniformément vers 0 sur le **segment** K, on peut trouver un rang N à partir duquel les intégrales sur K des fonction $|f - f_n|$ sont majorées par $\frac{\varepsilon}{2}$. On a alors clairement

$$\forall n \geqslant N \quad \int_I |f - f_n| < \varepsilon$$

ce qui démontre le résultat.

• Dans le cas général, comme dans ce qui précède, il suffit de démontrer que

$$\forall K \text{ segment} \subset I \qquad \lim_{n \to +\infty} \int_K |f - f_n| = 0$$

Ceci est conséquence immédiate du lemme **12-1**.26 où l'on posera $g_n = |f - f_n|$ et où on remplacera φ par 2φ. ■

Bien sûr, si on omet l'hypothèse de domination, la conclusion du théorème n'est plus valable. On sait bien que, même sur un segment, la convergence simple est insuffisante pour assurer la convergence des intégrales.

12-3.3 Intégration terme à terme d'une série de fonctions

THÉORÈME 12-3.3 Soit $(f_n)_n$ une suite de fonctions intégrables $I \to \mathbb{C}$ telle que

$$\left\{ \begin{array}{ll} \textbf{1)} & \textbf{La série } \sum_{n=0}^{+\infty} f_n \textbf{ converge simplement sur } I \\ \textbf{2)} & \textbf{La somme } f \textbf{ de cette série est continue par morceaux sur } I \\ \textbf{3)} & \textbf{La série } \sum_{n=0}^{+\infty} \int_I |f_n| \textbf{ est convergente} \end{array} \right.$$

Alors f est intégrable sur I et vérifie

$$\|f\|_1 \leqslant \sum_{n=0}^{+\infty} \|f_n\|_1 \qquad \textbf{et} \qquad \int_I f = \int_I \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_I f_n$$

Démonstration : la suite de fonctions

$$g_n = \inf \left(|f| , \sum_{k=0}^{n} |f_k| \right)$$

est une suite de fonctions continues par morceaux, intégrables, dont il est aisé de voir qu'elle converge simplement en croissant vers la fonction $|f|$. Comme pour tout n on a

$$\int_I g_n \leqslant \int_I \sum_{k=0}^{n} |f_k| \leqslant \sum_{k=0}^{+\infty} \int_I |f_k|$$

le théorème de convergence monotone assure l'intégrabilité de $|f|$ (donc celle de f) et on a

$$\int_I |f| = \lim_{n\to+\infty} \int_I g_n \leqslant \sum_{k=0}^{+\infty} \int_I |f_k|$$

De la même manière, en travaillant cette fois avec la série $\sum_{k=n+1}^{+\infty} f_k$, on obtiendra la majoration

$$\int_I \left| f - \sum_{k=0}^{n} f_k \right| \leqslant \sum_{k=n+1}^{+\infty} \int_I |f_k| \underset{n\to+\infty}{\longrightarrow} 0$$

ce qui démontre la convergence en moyenne de la série vers f, et en particulier la convergence des intégrales

$$\int_I f = \lim_{n\to+\infty} \sum_{k=0}^{n} \int_I f_n = \sum_{n=0}^{+\infty} \int_I f_n \quad \blacksquare$$

L'hypothèse $\sum_{n=0}^{+\infty} \|f_n\|_1$ convergente est importante, comme le montre un des exercices suivants.

EXERCICE 12-3.4 Montrer que

$$\int_0^{+\infty} \frac{\sin ax}{e^x - 1} dx = \sum_{n=1}^{+\infty} \frac{a}{n^2 + a^2}$$

EXERCICE 12-3.5 Si $f_n(x) = e^{-nx} - 2e^{-2nx}$, comparer

$$\int_0^{+\infty} \sum_{n=1}^{+\infty} f_n(x)\, dx \quad \text{et} \quad \sum_{n=1}^{+\infty} \int_0^{+\infty} f_n(x)\, dx$$

Que remarque-t-on ? Expliquer ce résultat.

EXERCICE 12-3.6 Pour quelles valeurs du complexe s la fonction

$$\Gamma(s) = \int_0^{+\infty} t^{s-1} e^{-t} dt$$

est-elle définie ? Montrer que Γ est développable en série entière au voisinage de 1. Dans quel domaine cette série entière représente-t-elle la fonction Γ ?

12-3.4 Intégrales dépendant d'un paramètre

Il s'agit ici de paraphraser le théorème de convergence dominée pour obtenir des théorèmes concernant les intégrales dépendant d'un paramètre.

12-3.4.1 Continuité

THÉORÈME 12-3.7 Soit A une partie d'un espace normé et I un intervalle de $\mathbb{R}$. On se donne une application

$$f : A \times I \to \mathbb{C}$$

telle que

$$\begin{cases} \textbf{1)} & \forall x \in A \quad f(x, \bullet) \textbf{ est continue par morceaux sur } I \\ \textbf{2)} & \exists \varphi \textbf{ intégrable positive sur } I \textbf{ telle que } \forall x \in A \ |f(x, \bullet)| \leqslant \varphi \textbf{ sur } I \end{cases}$$

La fonction

$$F(x) = \int_I f(x,t)\,dt$$

est alors définie sur A, et s'il existe $x_0 \in A$ tel que

$$\forall t \in I \quad x \mapsto f(x,t) \quad \textbf{est continue en } x_0$$

alors F est continue en x_0.

Démonstration : l'hypothèse de domination montre que F est bien définie sur A. Si $(u_n)_{n\in\mathbb{N}}$ est une suite de points de A qui converge vers x_0, la suite de fonctions

$$f_n = f(u_n, \bullet)$$

converge simplement vers la fonction $f(x_0, \bullet)$ et vérifie les hypothèses du théorème de convergence dominée. On a donc

$$\lim_{n\to+\infty} F(u_n) = \lim_{n\to+\infty} \int_I f(u_n,t)\,dt = \int_I f(x_0,t)\,dt = F(x_0)$$

ce qui prouve exactement la continuité de F en x_0. ∎

EXERCICE 12-3.8 Définition et continuité de la fonction

$$F(x) = \int_0^{+\infty} \frac{e^{-x(1+t^2)}}{1+t^2} \sin(xt)\,dt$$

12-3.4.2 Dérivabilité

THÉORÈME 12-3.9 On suppose que A est un intervalle de $\mathbb{R}$ et que f : $A \times I \to \mathbb{C}$ vérifie

$$\left\{\begin{array}{l} \textbf{1)}\ \forall x \in A \quad f(x,\bullet) \textbf{ est intégrable sur } I \\ \textbf{2)}\ f \textbf{ possède en tout point } (x,t) \textbf{ de } A\times I \textbf{ une dérivée partielle } \dfrac{\partial f}{\partial x}(x,t) \\ \textbf{3)}\ \forall x \in A \quad \dfrac{\partial f}{\partial x}(x,\bullet) \textbf{ est continue par morceaux sur } I \\ \textbf{4)}\ \textbf{Il existe } \varphi \textbf{ intégrable positive sur } I \textbf{ telle que} \\ \qquad \forall (x,t) \in A \times I \quad \left|\dfrac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t) \end{array}\right.$$

La fonction $F(x) = \int_I f(x,t)\,dt$ est alors dérivable sur A et

$$\forall x \in A \quad F'(x) = \int_I \frac{\partial f}{\partial x}(x,t)\,dt$$

Démonstration : Pour $x_0 \in A$ et $(h_n)_{n\in\mathbb{N}}$ suite de $\mathbb{R}^*$ tendant vers 0 telle que $x_0 + h_n \in A$, on a

$$\frac{F(x_0+h_n) - F(x_0)}{h_n} = \int_I f_n(t)\,dt$$

avec $f_n(t) = \dfrac{f(x_0+h_n,t) - f(x_0,t)}{h_n}$. Par définition de la dérivée partielle, la suite de fonctions f_n converge simplement sur I vers la fonction $t \mapsto \dfrac{\partial f}{\partial x}(x_0,t)$. L'inégalité des accroissements finis donne immédiatement

$$\forall n \quad \forall t \in I \quad |f_n(t)| \leqslant \varphi(t)$$

Le théorème de convergence dominée nous donne alors

$$\lim_{n\to+\infty} \frac{F(x_0+h_n) - F(x_0)}{h_n} = \int_I \frac{\partial f}{\partial x}(x_0,t)\,dt$$

Comme la suite h_n est quelconque, on a prouvé la dérivabilité de F en x_0. ■

REMARQUE 12-3.10 Pour être assuré de la dérivabilité de F en x_0, il suffit d'une domination de la dérivée partielle sur un ensemble de la forme $V_0 \times I$, où V_0 est un voisinage de x_0 (relatif à A).

COROLLAIRE 12-3.11 En particulier, avec les hypothèse précédentes, si $(x,t) \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur $A \times I$, la fonction F est de classe C^1 sur A.

EXERCICE 12-3.12 On pose

$$F(x) = \int_0^{+\infty} \frac{e^{-x(1+t^2)}}{1+t^2} dt$$

Définition, continuité et dérivabilité de F. Montrer que, pour $x \geqslant 0$

$$\frac{\pi}{2} - F(x) = \frac{1}{2} \int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt \int_0^{x} \frac{e^{-t}}{\sqrt{t}} dt$$

En déduire que

$$\int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt = \sqrt{\pi}$$

12-3.5 La fonction Γ

Les théorèmes précédents permettent de démontrer les principales propriétés de la fonction Γ.

- La fonction $s \mapsto \Gamma(s) = \displaystyle\int_0^{+\infty} t^{s-1} e^{-t} dt$ est définie et continue dans le demi-plan
$$\mathcal{P}^+ = \{s \in \mathbb{C} \mid \operatorname{Re} s > 0\}$$
- Sa restriction à $]0, +\infty[$ est $\mathcal{C}^\infty$ et
$$\forall k \in \mathbb{N}^* \quad \forall s > 0 \quad \Gamma^{(k)}(s) = \int_0^{+\infty} (\ln t)^k t^{s-1} e^{-t} \, dt$$
- $\forall s \in \mathcal{P}^+ \quad \Gamma(s+1) = s\,\Gamma(s)$ (intégration par parties). En particulier, pour n entier non nul, $\Gamma(n) = (n-1)!$
- $\Gamma(s) \underset{0^+}{\sim} \dfrac{1}{s}$
- L'exercice **12-3**.12 montre que $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$.

EXERCICE 12-3.13 Montrer que la fonction $s \mapsto \ln \Gamma(s)$ est convexe sur $]0, +\infty[$

EXERCICE 12-3.14 Montrer que, pour $s \in \mathcal{P}^+$

$$\Gamma(s) = \lim_{n \to +\infty} \int_0^n t^{s-1} \left(1 - \frac{t}{n}\right)^n dt$$

En déduire que, pour $s \in \mathcal{P}^+$,

$$\Gamma(s) = \lim_{n \to +\infty} \frac{n^s n!}{s(s+1) \cdots (s+n)}$$

Montrer que la limite existe pour tout s complexe différent d'un entier négatif.

12-4 Intégrales impropres

Nous avons déjà vu (*cf.* exemple **12-2**.10) que lorsqu'une fonction continue par morceaux sur un intervalle $[a,b[\to \mathbb{C}$ vérifie

$$\lim_{x\to b^-}\int_a^x f(t)\,dt \quad \text{existe}$$

la fonction f n'est pas forcément intégrable sur $[a,b[$ (sauf bien sûr si f est réelle et garde un signe constant au voisinage de b). Le contre-exemple $f(t)=\dfrac{\sin t}{t}$ a été donné sur $[0,+\infty[$.

DÉFINITION 12-4.1 *Si f est une fonction continue par morceaux sur un intervalle $[a,b[\to\mathbb{C}$, on dit que l'intégrale*

$$\int_a^b f(t)\,dt \quad \textit{existe de manière impropre}$$

ssi $\lim\limits_{x\to b^-}\int_a^x f(t)\,dt$ *existe sans que f soit intégrable sur $[a,b[$. On note alors,* ***par abus d'écriture****,*

$$\int_a^b f(t)\,dt = \lim_{x\to b^-}\int_a^x f(t)\,dt$$

Pour éviter cet abus de notation, lorsqu'il s'agit d'être plus précis, on utilisera aussi la notation $\displaystyle\int_a^{\to b} f(t)\,dt$.

On aurait une définition analogue sur un intervalle de la forme $]a,b]$. Pour continuer l'analogie avec la théorie des séries, disons que la notion de fonction intégrable correspond à la notion de série absolument convergente, alors que celle d'intégrale impropre correspond à la notion de série semi-convergente.

Pour étudier l'existence de $\lim\limits_{x\to b^-}\int_a^x f(t)\,dt$, la démarche "raisonnable" est d'abord de tenter de prouver l'intégrabilité de f, et si cela échoue (soit parce que f n'est pas intégrable, soit parce que ... on ne sait pas faire) essayer d'autres techniques adaptées aux intégrales impropres. Ce sont essentiellement :

- Une intégration par parties, pour augmenter le degré d'un dénominateur le plus souvent. On peut prouver ainsi par exemple l'existence de l'intégrale $\displaystyle\int_1^{\to+\infty}\frac{\cos t}{\sqrt{t^2+1}}dt$ (on peut dire avant toute chose que si l'intégrale existe, c'est forcément une intégrale impropre, puisque $\left|\dfrac{\cos t}{\sqrt{t^2+1}}\right| \underset{+\infty}{\sim} \dfrac{|\cos t|}{t}$ et on sait que cette fonction n'est pas intégrable).

- L'utilisation d'une "décomposition" : si $f(t)\underset{b^-}{\sim} g(t)$ et si $\displaystyle\int_a^{\to b} g(t)\,dt$ existe, l'existence de l'intégrale (forcément impropre -pourquoi ?-) de f sera subordonnée à celle de l'intégrale de $f-g$. On pourrait traiter de cette manière aussi $\displaystyle\int_1^{\to+\infty}\frac{\cos t}{\sqrt{t^2+1}}dt$. Voir aussi l'exercice suivant.

- L'utilisation d'une série : pour prouver l'existence de $\int_1^{\to+\infty} \frac{\cos t}{\sqrt{t^2+1}}dt$, on peut d'abord montrer la convergence de la série de terme général

$$u_n = \int_{\frac{\pi}{2}+n\pi}^{\frac{\pi}{2}+(n+1)\pi} \frac{\cos t}{\sqrt{t^2+1}}dt$$

(la réponse à la question -pourquoi ces bornes d'intégration ?- contient déjà la quasi-totalité de la solution). Puis, si x est "grand" et vérifie un encadrement

$$\frac{\pi}{2} + n\pi \leqslant x < \frac{\pi}{2} + (n+1)\pi$$

par quoi peut-on majorer $\left| \int_{\frac{\pi}{2}+n\pi}^{x} \frac{\cos t}{\sqrt{t^2+1}}dt \right|$?

EXERCICE 12-4.2 Discuter, suivant les valeurs de $\alpha > 0$, l'existence de l'intégrale

$$\int_1^{+\infty} \ln\left(1 + \frac{\sin t}{t^\alpha}\right) dt$$

Solution : la fonction $t \mapsto f(t) = \ln\left(1 + \frac{\sin t}{t^\alpha}\right)$ est continue sur $[1, +\infty[$, car pour $t \geqslant 1$ on a $\left|\frac{\sin t}{t^\alpha}\right| < 1$. Au voisinage de l'infini, $f(t) \sim \frac{\sin t}{t^\alpha}$, **donc** $|f(t)| \sim \frac{|\sin t|}{t^\alpha}$. Cette dernière fonction est intégrable ssi $\alpha > 1$. On est donc assuré de l'existence de l'intégrale pour ces valeurs de α. Si $\alpha \in]0, 1]$, l'intégrale ne peut être qu'impropre. Au voisinage de $+\infty$, on a

$$f(t) = \frac{\sin t}{t^\alpha} - \frac{1}{2}\frac{\sin^2 t}{t^{2\alpha}} + o\left(\frac{\sin^2 t}{t^{2\alpha}}\right)$$

ce qui peut s'écrire aussi

$$f(t) = \frac{\sin t}{t^\alpha} - g(t)$$

avec $g(t) \underset{+\infty}{\sim} \frac{1}{2}\frac{\sin^2 t}{t^{2\alpha}}$ qui est donc **positive au voisinage de** $+\infty$. Comme $\int_1^{+\infty} \frac{\sin t}{t^\alpha}dt$ existe, on est ramené à l'existence de l'intégrale de g. Il est aisé de voir que g est intégrable ssi $\alpha > \frac{1}{2}$.

Attention ! L'exercice précédent montre que deux fonctions équivalentes en b peuvent avoir l'une une intégrale impropre qui existe et l'autre une intégrale qui n'existe pas. C'est le cas par exemple pour $\alpha \in \left]0, \frac{1}{2}\right]$ où $\int_1^{\to+\infty} \frac{\sin t}{t^\alpha}dt$ existe (intégration par parties). On rencontre ici la même difficulté que pour les séries semi-convergentes.

REMARQUE 12-4.3 L'étude d'intégrales impropres dépendant d'un paramètre n'est pas possible à l'aide des théorèmes généraux qui précèdent, qui supposent tous l'intégrabilité des fonctions manipulées. L'étude doit donc se faire au cas par cas. On retiendra une méthode de transformation (intégration par parties le plus souvent) qui permet de se ramener à des fonctions intégrables. On peut aussi, pour étudier

$$F(x) = \int_a^{\to b} f(x,t)\,dt$$

appliquer les théorèmes généraux aux fonctions

$$F_n(x) = \int_a^{a_n} f(x,t)\,dt$$

où a_n est une suite croissante de points de $[a, b[$ qui converge vers b, et essayer d'obtenir des résultats de régularité de leur limite simple F par des arguments de convergence uniforme.

12-5 Relation série-intégrale

L'utilisation de la notion d'intégrale est très fréquente pour étudier les séries. Rappelons le théorème suivant

12-5.1 Cas de fonctions positives

THÉORÈME 12-5.1 Si $f : [0, +\infty[\to \mathbb{R}^+$ est décroissante et continue par morceaux, la série de terme général (positif)

$$w_n = \int_{n-1}^{n} f(t)\,dt - f(n)$$

est convergente.

Rappelons que, lorsque f **est intégrable sur** $[0, +\infty[$, on déduit la convergence de la série de terme général $u_n = f(n)$ et l'encadrement du reste d'ordre n (faire un dessin !).

$$\int_{n+1}^{+\infty} f(t)\,dt \leqslant \sum_{k=n+1}^{+\infty} f(k) \leqslant \int_{n}^{+\infty} f(t)\,dt$$

Lorsque f n'est pas intégrable, la série est divergente. La somme partielle

$$S_n = f(1) + \cdots + f(n)$$

et la suite $F(n) = \displaystyle\int_0^n f(t)\,dt$ forment (pour $n \to +\infty$) des infiniment grands "parallèles". Il existe un réel A tel que

$$f(1) + \cdots + f(n) \underset{n\to+\infty}{=} \int_0^n f(t)\,dt + A + o(1)$$

12-5.2 Cas de fonctions complexes

Lorsque $f : [0, +\infty[\to \mathbb{C}$ est continue par morceaux, on ne pourra plus procéder par encadrements comme dans la démonstration du théorème précédent. Cependant, lorsque la fonction f varie "assez peu" au voisinage de l'infini, une "bonne" approximation de $u_n = f(n)$ sera encore obtenue par $v_n = \int_{n-1}^{n} f(t)\, dt$, l'intérêt étant que les sommes partielles de la série de terme général v_n s'écrivent très simplement par la relation de Chasles. On retiendra en particulier le "théorème" suivant

THÉORÈME 12-5.2 Si $f : [0, +\infty[\to \mathbb{C}$ est de classe C^1, avec f' intégrable sur $[0, +\infty[$, alors la série de terme général

$$w_n = \int_{n-1}^{n} f(t)\, dt - f(n)$$

est absolument convergente.

Démonstration : Une intégration par parties donne

$$w_n = \int_{n-1}^{n} (t - (n-1))\, f'(t)\, dt$$

(c'est aussi, si on veut, la formule de Taylor avec reste sous forme d'intégrale, pour évaluer $F(n-1) - F(n)$ avec $F(x) = \int_a^x f(t)\, dt$). On a alors la majoration

$$|w_n| \leqslant \int_{n-1}^{n} |f'(t)|\, dt$$

qui permet de conclure, puisque l'intégrabilité de f' entraîne la convergence de la série majorante. ■

EXERCICE 12-5.3 Fonction ζ de Riemann : Montrer que le domaine de définition de la fonction d'une variable complexe

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$$

est exactement le demi-plan $\{s \in \mathbb{C} \mid \operatorname{Re} s > 1\}$. Montrer que, pour s réel tendant vers 1 par valeurs supérieures on a

$$\zeta(s) \sim \frac{1}{s-1}$$

Solution : Comme $\left|\frac{1}{n^s}\right| = \frac{1}{n^{\operatorname{Re} s}}$, la série est absolument convergente pour tout s tel que $\operatorname{Re} s > 1$. Elle est évidemment grossièrement divergente pour $\operatorname{Re} s \leqslant 0$. Il reste donc à prouver la divergence pour $\operatorname{Re} s \in]0, 1]$. La fonction $t \mapsto t^{-s}$ est évidemment C^∞ sur $[1, +\infty[$,

sa dérivée vérifie $|f'(t)| = \frac{|s|}{t^{\mathrm{Re}\, s+1}}$ et est intégrable sur $[1, +\infty[$. Le théorème précédent s'applique. Le cas $s = 1$ est trivial. Sinon

$$\int_{n-1}^{n} \frac{dt}{t^s} = \frac{1}{1-s}\left[\frac{1}{n^{s-1}} - \frac{1}{(n-1)^{s-1}}\right]$$

est le terme général de la série télescopique associée à la suite divergente $\frac{1}{n^{s-1}}$ (le cas $\mathrm{Re}\, s = 1$ requiert peut-être une attention particulière). Pour trouver l'équivalent de $\zeta(s)$ en 1^+, il suffit d'encadrer la somme de la série par deux intégrales. ■

EXERCICE 12-5.4 Discuter, en fonction de α, la nature de la série de terme général

$$u_n = \frac{\sin \pi\sqrt{n}}{n^\alpha}$$

(Dans le cas où le théorème précédent ne permet pas de conclure, on doit pouvoir prouver la divergence en niant le critère de Cauchy).

12-5.3 Utilisation de la relation de Chasles

Les théorèmes qui précèdent permettent souvent de ramener l'étude d'une série à celle d'une intégrale. Inversement, l'utilisation d'une série permet parfois de prouver l'existence d'une intégrale (impropre ou non). Nous étudierons le cas $I = [a, +\infty[$ avec f continue par morceaux, le cas général $[a, b[$ se ramenant à des études analogues.

- Si f **est positive** et s'il existe une suite a_n croissante tendant vers $+\infty$ telle que la série de terme général

$$u_n = \int_{a_n}^{a_{n+1}} f(t)\, dt$$

converge, alors f **est intégrable** et on a (évidemment)

$$\sum_{n=0}^{+\infty} u_n = \int_{a_0}^{+\infty} f(t)\, dt$$

- Si f **est complexe**, l'existence d'une telle suite ne peut évidemment pas prouver l'intégrabilité de f. Le cas $f(t) = \sin t$ et $a_n = 2n\pi$ montre qu'elle n'entraîne pas non plus l'existence d'une intégrale impropre. On pourra cependant conclure (exercice) si

$$\lim_{n\to+\infty} \int_{a_n}^{a_{n+1}} |f(t)|\, dt = 0$$

C'est la méthode qui a été suggérée pour prouver l'existence de $\int_1^{+\infty} \frac{\sin t}{t} dt$ en utilisant la suite $a_n = n\pi$ au début de la section sur les intégrales impropres.

EXERCICE 12-5.5 Etudier l'intégrabilité sur $[0,+\infty[$ de la fonction

$$f(t) = \frac{t}{1+t^6\sin^2 t}$$

EXERCICE 12-5.6 Même question avec

$$g(t) = \frac{e^t}{e^{-t}+e^{2t}\,|\sin t|}$$

12-6 Compléments : transformations de Fourier et Laplace

Sous forme d'exercices, on illustre ici l'utilisation des théorèmes relatifs aux intégrales à paramètres pour obtenir quelques résultats élémentaires sur les transformées de Fourier et Laplace. En particulier, on obtient (sous certaines hypothèses) des résultats d'injectivité, bien utiles dans la pratique : toute l'information sur une fonction f est contenue dans sa transformée de Fourier (ou de Laplace).

12-6.1 Transformation de Fourier

DÉFINITION 12-6.1 *Soit $f : \mathbb{R} \to \mathbb{C}$ une fonction continue par morceaux et intégrable. Sa transformée de Fourier est la fonction $\widehat{f}$ définie sur $\mathbb{R}$ par*

$$\forall x \in \mathbb{R} \quad \widehat{f}(x) = \int_{\mathbb{R}} f(t)\, e^{-2i\pi xt} dt$$

Cette fonction est parfaitement définie, puisque $t \mapsto f(t)\,e^{-2i\pi xt}$ est continue par morceaux sur $\mathbb{R}$, avec $|f(t)\,e^{-2i\pi xt}| = |f(t)|$. Dans tout ce qui suit, f représente une fonction intégrable sur $\mathbb{R}$.

EXERCICE 12-6.2 Montrer que $\widehat{f}$ est une fonction continue bornée sur $\mathbb{R}$, vérifiant

$$\left\|\widehat{f}\right\|_\infty \leqslant \|f\|_1$$

En déduire que la transformation de Fourier $f \mapsto \widehat{f}$ est une application linéaire continue de $\left(\mathcal{L}^1(\mathbb{R},\mathbb{C}), \|\,\|_1\right)$ dans l'espace $\left(\mathcal{C}_b^0(\mathbb{R},\mathbb{C}), \|\,\|_\infty\right)$ des fonctions continues bornées sur $\mathbb{R}$. Si g est une fonction en escalier (nulle hors d'un segment $[-A, A]$), montrer que

$$\lim_{x\to\pm\infty} \widehat{g}(x) = 0$$

Si f est intégrable, montrer qu'on peut trouver une suite $(g_n)_{n\in\mathbb{N}}$ de fonctions en escalier telles que $\lim\limits_{n\to+\infty} \|f-g_n\|_1 = 0$. En déduire que $\widehat{f}$ tend vers 0 en $\pm\infty$.

EXERCICE 12-6.3 Transformation de Fourier et dérivation :

On suppose que f est continue et C^1 par morceaux, avec f' intégrable sur $\mathbb{R}$. Montrer que

$$\lim_{\pm\infty} f = 0$$

et que la transformée de Fourier de sa dérivée vérifie

$$\forall x \in \mathbb{R} \quad \widehat{f'}(x) = 2i\pi x \, \widehat{f}(x)$$

La dérivation se traduit donc, au niveau de la transformée de Fourier, par la multiplication par un monôme.

A l'inverse, si on suppose seulement f intégrable et $g : t \mapsto t\, f(t)$ intégrable, montrer que $\widehat{f}$ est de classe C^1 sur $\mathbb{R}$ avec

$$\forall x \in \mathbb{R} \quad \left(\widehat{f}\right)'(x) = -2i\pi \widehat{g}(x)$$

Que peut-on en déduire si on suppose $t \mapsto t^p f(t)$ intégrable (avec $p \geqslant 2$ entier naturel quelconque) ?

On note

$$\mathcal{S}(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}, \mathbb{C}) \mid \forall k, l \in \mathbb{N} \quad \lim_{x \to \pm\infty} x^k f^{(l)}(x) = 0 \right\}$$

Les fonctions de $\mathcal{S}(\mathbb{R})$ sont C^∞ et toutes leurs dérivées sont dites à "décroissance rapide" à l'infini. Montrer que $\mathcal{S}(\mathbb{R})$ est une sous-algèbre de $C^\infty(\mathbb{R}, \mathbb{C})$, que $\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^1(\mathbb{R}, \mathbb{C})$ et que

$$\forall f \in \mathcal{S}(\mathbb{R}) \quad \widehat{f} \in \mathcal{S}(\mathbb{R})$$

EXERCICE 12-6.4 Transformée de Fourier d'une gaussienne : soit φ définie sur $\mathbb{R}$ par

$$\forall t \in \mathbb{R} \quad \varphi(t) = e^{-\pi t^2}$$

Montrer que $\varphi \in \mathcal{S}(\mathbb{R})$, et que sa transformée de Fourier est solution de l'équation différentielle

$$y' + 2\pi x\, y = 0$$

En déduire que

$$\varphi = \widehat{\varphi}$$

(on rappelle que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$).

Plus généralement, si $\sigma > 0$, on pose

$$f_\sigma(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

Vérifier que $\int_{\mathbb{R}} f_\sigma = 1$, et, par un changement de variable linéaire, montrer que

$$\forall x \in \mathbb{R} \quad \widehat{f_\sigma}(x) = e^{-2\pi^2\sigma^2 x^2}$$

Sur cet exemple très simple, on voit apparaître une propriété importante de la transformation de Fourier : si σ est "très petit", la fonction f_σ est très "localisée" au voisinage de 0 (faire un dessin) et sa transformée de Fourier est alors très "étalée". Voir à ce sujet l'exercice **16-3**.16.

EXERCICE 12-6.5 Inversion de Fourier : soient f et g deux fonctions continues sur $\mathbb{R}$. Montrer que, pour A et $B > 0$ quelconques, on a

$$\int_{-A}^{A}\left(\int_{-B}^{B} g(t)\, e^{-2i\pi xt} dt\right) f(x)\, dx = \int_{-B}^{B}\left(\int_{-A}^{A} f(x)\, e^{-2i\pi xt} dx\right) g(t)\, dt$$

En déduire que, si f, g, $\widehat{f}$ et $\widehat{g}$ sont intégrables sur $\mathbb{R}$, on a

$$\int_{\mathbb{R}} f(t)\, \widehat{g}(t)\, dt = \int_{\mathbb{R}} g(t)\, \widehat{f}(t)\, dt$$

En prenant pour g la fonction $\widehat{f_\sigma}$ étudiée précédemment, et en faisant tendre σ vers 0, montrer que

$$f \text{ continue et } \widehat{f} \text{ intégrable} \Rightarrow f(0) = \int_{\mathbb{R}} \widehat{f}(x)\, dx$$

En considérant enfin la fonction $x \mapsto f(x+t)$, en déduire la **formule d'inversion de Fourier**

$$f \text{ continue et } \widehat{f} \text{ intégrable} \Rightarrow \forall t \in \mathbb{R} \quad f(t) = \int_{\mathbb{R}} \widehat{f}(x)\, e^{2i\pi tx}\, dx$$

EXERCICE 12-6.6 Déduire des exercices qui précèdent le fait que la transformation de Fourier induit un automorphisme sur l'espace vectoriel $\mathcal{S}(\mathbb{R})$. Montrer aussi que

$$\forall f \in \mathcal{S}(\mathbb{R}) \quad \int_{\mathbb{R}} |f|^2 = \int_{\mathbb{R}} |\widehat{f}|^2$$

12-6.2 Transformation de Laplace

Dans cette section, nous développerons des techniques analogues à celles qui précèdent pour prouver l'injectivité de la transformation de Laplace :

EXERCICE 12-6.7 Soit $f : [0, +\infty[\to \mathbb{R}$ continue et bornée. On définit sa transformée de Laplace par

$$F(p) = \int_0^{+\infty} f(t)\, e^{-pt}\, dt$$

lorsque cette intégrale existe. On posera de même

$$G(p) = \int_0^{+\infty} t f(t)\, e^{-pt}\, dt$$

- Montrer que $F(p)$ et $G(p)$ sont définis pour tout $p > 0$. Déterminer leurs limites en $+\infty$.
- Montrer que F est dérivable sur $]0, +\infty[$ avec

$$\forall p > 0 \quad F'(p) = -G(p)$$

  En déduire que F est indéfiniment dérivable sur $\mathbb{R}^{*+}$.
- Si f est de plus supposée de classe $\mathcal{C}^1$ sur $[0, +\infty[$, avec f' bornée, quel lien y a-t-il entre la transformée de Laplace de f et celle de f' ?

- Si f est intégrable sur $[0,+\infty[$, montrer que F est continue en 0. Calculer la transformée de Laplace de $t \mapsto e^{-t^2}$.

EXERCICE 12-6.8 On définit la fonction $h : \mathbb{R} \to \mathbb{R}$ par

$$h(t) = e^t e^{-e^t} = \sum_{k=0}^{+\infty} \frac{(-1)^k}{k!} e^{(k+1)t}$$

- Montrer que h est intégrable sur $\mathbb{R}$ et que $\int_{\mathbb{R}} h = 1$.
- Si n est un entier naturel non nul, on définit $h_n : \mathbb{R} \to \mathbb{R}$ par $h_n(t) = n\,h(nt)$. Montrer que

$$\forall a > 0 \quad \lim_{n \to +\infty} \int_{-a}^{a} h_n(t)\; dt = 1$$

- En déduire que, pour f continue bornée : $[0,+\infty[\to \mathbb{R}$ et $x > 0$

$$\lim_{n \to +\infty} \int_0^{+\infty} h_n(x-t)\; f(t)\; dt = f(x)$$

(après avoir vérifié l'existence de ces intégrales).

On suppose toujours que f est continue bornée et que F est sa transformée de Laplace.

- Montrer que, pour tout $x > 0$ et tout $n \in \mathbb{N}^*$

$$n \sum_{k=0}^{+\infty} \frac{(-1)^k}{k!} e^{n(k+1)x} F(n(k+1)) = \int_0^{+\infty} h_n(x-t)\; f(t)\; dt$$

- En déduire que, s'il existe $A > 0$ tel que F soit identiquement nulle sur $[A,+\infty[$, alors f est nulle. Que peut-on dire de deux fonctions continues bornées ayant mêmes transformées de Laplace ?

12-6.3 Produit de convolution : un procédé de régularisation

Si on reprend les deux études qui précèdent, on s'aperçoit qu'elles utilisent une même technique, basée sur la notion de produit de convolution. Si $f : \mathbb{R} \to \mathbb{C}$ est une fonction intégrable et si $g : \mathbb{R} \to \mathbb{C}$ est continue bornée[1], on définit le produit de convolution de f par g :

$$\forall x \in \mathbb{R} \quad f * g(x) = \int_{\mathbb{R}} f(x-t)\; g(t)\; dt$$

[1] Les hypothèses faites ici sur f et g assurent l'existence de l'intégrale définissant le produit de convolution. Ce dernier pourrait être défini dans d'autres situations, par exemple en supposant f et g de carrés intégrables.

Par changement de variable, on a aussi[2]

$$f * g(x) = \int_{\mathbb{R}} g(x-t)\, f(t)\, dt$$

En utilisant une de ces deux expressions, on pourra souvent montrer que la régularité de $f * g$ est au moins celle de f ou g. Par exemple :

EXERCICE 12-6.9 Si f est intégrable sur $\mathbb{R}$ et g est une fonction C^∞ dont toutes les dérivées sont bornées, montrer que $f * g$ est C^∞ sur $\mathbb{R}$ et que

$$\forall k \in \mathbb{N}^* \quad (f * g)^{(k)} = f * g^{(k)}$$

De plus, si on choisit bien la fonction g, $f * g$ sera une "bonne approximation" de f possédant la régularité de g : l'opération de convolution est "régularisante". Par exemple, on a :

EXERCICE 12-6.10 Soit h une fonction C^∞ à dérivées bornées, intégrable sur $\mathbb{R}$ avec

$$h \geqslant 0 \quad \text{et} \quad \int_{\mathbb{R}} h = 1$$

(on peut interpréter h comme une densité sur la droite réelle, la masse totale répartie étant égale à 1 ; on peut par exemple prendre $h(x) = e^{-\pi x^2}$). Si on prend

$$h_n(t) = n\, h(nt)$$

h_n correspond encore à une répartition de masse totale égale à 1, mais qui se concentre au voisinage de l'origine :

$$\forall a > 0 \quad \lim_{n \to +\infty} \int_{-a}^{a} h_n(t)\, dt = 1$$

Si f est une fonction intégrable bornée sur $\mathbb{R}$, la suite $(f_n)_{n\in\mathbb{N}} = (f * h_n)_{n\in\mathbb{N}}$ est formée de fonctions C^∞ (d'après l'exercice qui précède). Montrer que, si x_0 est un point de continuité de f, on a

$$\lim_{n \to +\infty} f_n(x_0) = f(x_0)$$

Si de plus, f est continue sur un intervalle $]a, b[$, montrer que la convergence est uniforme sur tout segment inclus dans $]a, b[$.

Nous reverrons ultérieurement ce genre de technique (voir notamment la section **16-2.4.2**). Terminons par une utilisation de la convolution pour démontrer le théorème de Weierstrass :

[2]Si g est de plus nulle hors d'un segment $[-A, A]$ (on dit alors que g est à support compact), on aura aussi

$$f * g(x) = \int_{-A}^{A} f(x-t)\, g(t)\, dt = \int_{x-A}^{x+A} f(t)\, g(x-t)\, dt$$

EXERCICE 12-6.11 Si $n \in \mathbb{N}^*$, on définit $f_n : \mathbb{R} \to \mathbb{R}$ par $f_n(t) = 0$ si $|t| \geqslant 1$ et

$$f_n(t) = \frac{1}{c_n}\left(1 - t^2\right)^n \quad \text{si} \quad |t| \leqslant 1$$

où c_n est choisi pour avoir

$$\int_{\mathbb{R}} f_n = \int_{-1}^{1} f_n(t)\, dt = 1$$

Montrer que

$$c_n \geqslant 2\int_0^1 (1-t)^n\, dt$$

et en déduire que

$$\forall a > 0 \quad \lim_{n \to +\infty} \int_{-a}^{a} f_n(t)\, dt = 1$$

Soit f une fonction continue sur $\left[-\frac{1}{4}, \frac{1}{4}\right]$ à valeurs dans $\mathbb{C}$. On prolonge f en une fonction g continue, nulle hors de $\left[-\frac{1}{2}, \frac{1}{2}\right]$ et affine sur chacun des segments $\left[-\frac{1}{2}, -\frac{1}{4}\right]$ et $\left[\frac{1}{4}, \frac{1}{2}\right]$. Montrer que la suite $(f_n * g)$ converge uniformément vers f sur $\left[-\frac{1}{4}, \frac{1}{4}\right]$. Pour $|x| \leqslant \frac{1}{4}$, montrer que

$$f_n * g(x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_n(x-t)\, g(t)\, dt = \frac{1}{c_n}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(1 - (x-t)^2\right)^n g(t)\, dt$$

En déduire que toute fonction continue sur un segment est limite uniforme d'une suite de polynômes.

12-7 Exercices

EXERCICE 12-7.1 Soit $f \in \mathcal{C}^2(\mathbb{R}^+, \mathbb{R})$ telle que f^2 et $(f'')^2$ soient intégrables. Montrer que $\int_0^{+\infty} (f')^2$ existe.

EXERCICE 12-7.2 Soit $f \in \mathcal{C}^0(\mathbb{R}, \mathbb{R})$ telle que $\int_{-\infty}^{+\infty} |f(x+t) - f(x)|dx$ existe pour tout $t \in \mathbb{R}$ et soit $o(t)$ au voisinage de 0. Que peut-on dire de f ?

EXERCICE 12-7.3 Existence et calcul de $\int_0^{+\infty} e^{-(t^2 + \frac{1}{t^2})} dt$.

EXERCICE 12-7.4 Soit f continue par morceaux, positive décroissante telle que l'intégrale $\int_0^{+\infty} f(t)dt$ existe. Déterminer

$$\lim_{h \to 0^+} \sum_{n=0}^{+\infty} h f(nh)$$

Application : si $\alpha > -1$, donner un équivalent, pour x tendant vers 1 par valeurs inférieures, de

$$\sum_{n=1}^{+\infty} n^{\alpha} x^{n}$$

(on pourra poser $h = -\ln x$).

EXERCICE 12-7.5 On suppose f continue de $[0, +\infty[$ dans $\mathbb{R}$ et de carré intégrable sur $\mathbb{R}^+$. On pose $F(x) = \int_0^x f(t)dt$. Montrer que, pour tout x on a

$$\int_0^x \frac{F(t)^2}{t^2} dt \leqslant 2 \int_0^x \frac{F(t)f(t)}{t} dt.$$

En déduire que

$$\int_0^{+\infty} \frac{F(t)^2}{t^2} dt \leqslant 4 \int_0^{+\infty} f^2(t) dt.$$

Traiter le cas de l'égalité.

EXERCICE 12-7.6 Soit $f \in C^1([1, +\infty[, \mathbb{R}_+^*)$ avec $\frac{f'}{f} \sim \frac{\mu}{x}$ en $+\infty$ avec $\mu \in]-1, +\infty[$ et $\mu \neq 0$. Montrer que f n'est pas intégrable sur $[1, +\infty[$ et que $\int_1^x f(t)dt \sim \frac{xf(x)}{\mu + 1}$ en $+\infty$. Trouver un équivalent en $+\infty$ de $\int_1^x \sqrt{t} e^{-\frac{1}{t}} dt$.

EXERCICE 12-7.7 Soit f continue : $]0, +\infty[\to \mathbb{C}$. On suppose que $\frac{f(x)}{x}$ a une limite en 0 et en $+\infty$. Soient $a > 0$ et $b > 0$, étudier l'existence et la valeur de

$$\int_0^{\to +\infty} \frac{bf(ax) - af(bx)}{x^2} dx$$

EXERCICE 12-7.8 Soit λ_n une suite de réels tendant vers $+\infty$ et $f : \mathbb{R}_+ \to \mathbb{R}$ continue par morceaux et intégrable sur $\mathbb{R}$. Montrer que

$$\lim_{n \to +\infty} \int_0^{+\infty} f(t) \sin \lambda_n t dt = 0$$

EXERCICE 12-7.9 Calculer $\lim_{n \to +\infty} \sum_{k=0}^{n-1} \frac{1}{\sqrt{n^2 - k^2}}$.

EXERCICE 12-7.10 Existence et calcul de $I = \int_0^1 \ln(1 - x^2) \frac{dx}{x^2}$.

EXERCICE 12-7.11 Discuter, selon les valeurs du réel α, l'existence de l'intégrale (éventuellement impropre) $\int_0^{+\infty} \ln\left(1 + \frac{\sin x}{x^{\alpha}}\right) dx$.

EXERCICE 12-7.12 Montrer que, pour $a > 1$ on a $\int_a^x \frac{e^{\sqrt{\ln t}}}{t \ln t} dt \sim \frac{2e^{\sqrt{\ln x}}}{\ln x}$ pour $x \to \infty$.

EXERCICE 12-7.13 Discuter en fonction du paramètre réel α le comportement de la suite

$$I_n(\alpha) = \int_0^1 e^{[\alpha(x+\frac{x^2}{2}+\cdots+\frac{x^n}{n})]}dx$$

EXERCICE 12-7.14 Calculer $I_n = \displaystyle\int_0^1 \frac{\ln(1+x+\cdots+x^n)}{x}dx$ (on utilisera des séries).

EXERCICE 12-7.15 Existence de $I_n = \displaystyle\int_0^{+\infty} \frac{dt}{(1+t^3)^n}$. Limite de I_n lorsque $n \to +\infty$. Convergence et somme de la série de terme général $(-1)^n I_n$.

EXERCICE 12-7.16 Existence et calcul de $\displaystyle\int_0^{+\infty} \frac{\operatorname{Arctan}\left(\frac{1}{x}\right)}{1+x^2}dx$.

EXERCICE 12-7.17 Existence de $\displaystyle\int_1^{+\infty} \frac{dx}{|x^\alpha - 1|^\beta}$ où α et $\beta \in \mathbb{R}$.

EXERCICE 12-7.18 On pose $f(x) = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{x+n}}$ pour $x > 0$. Etudier la continuité et la dérivabilité de f, ses limites en 0 et $+\infty$. Montrer que $f(x) \sim \dfrac{1}{2\sqrt{x}}$ en $+\infty$, puis établir l'existence d'un développement limité à tout ordre en $+\infty$ selon les puissances de $\dfrac{1}{x}$ de $\sqrt{x}f(x)$. Montrer que

$$f(x) = \int_0^{+\infty} \frac{e^{-tx}}{\sqrt{t}(1+e^{-t})}dt$$

EXERCICE 12-7.19 On pose $h(a) = \displaystyle\int_0^{+\infty} e^{-x^2}\cos(ax)dx$. Montrer que h est de classe C^1 sur $\mathbb{R}$ et la calculer.

EXERCICE 12-7.20 On pose $F(\alpha) = \displaystyle\int_0^{+\infty} \frac{\sin \alpha x}{x(x^2+1)}dx$. Existence de F, continuité, dérivabilité, calcul de F' et F''. En déduire la valeur de $F(\alpha)$ en fonction de $\displaystyle\int_0^{\to+\infty} \frac{\sin t}{t}dt$, puis la valeur de cette dernière intégrale.

EXERCICE 12-7.21 Comportement en $+\infty$ de $\phi(t) = \displaystyle\int_0^1 \frac{dx}{(1+x+x^2)^t}$. En donner un équivalent.

EXERCICE 12-7.22 f est continue sur $[0, +\infty[$ et $F(x) = \displaystyle\int_0^{+\infty} \frac{f(t)}{t+x}dt$

1. On suppose f intégrable. Etudier la définition et la continuité de F sur $]0, +\infty[$, puis $\displaystyle\lim_{x\to+\infty} xF(x)$.

2. Reprendre ces questions, en supposant seulement l'intégrale $\displaystyle\int_0^{\to+\infty} f(x)dx$ impropre.

3. On suppose que $\int_0^{\to+\infty} \frac{f(t)}{t}dt$ existe. Etudier $\lim_{x\to 0^+} F(x)$.

EXERCICE 12-7.23 Montrer que $x \mapsto \varphi(x) = \int_0^{+\infty} \frac{\sin xt}{t(1+t^4)}dt$ est $\mathcal{C}^3$ sur $\mathbb{R}$ et $\mathcal{C}^4$ sur $\mathbb{R}^{*+}$. Montrer que φ est solution sur $\mathbb{R}^{*+}$ d'une équation différentielle linéaire d'ordre 4 à coefficients constants. Déterminer $\varphi(x)$ pour $x > 0$.

EXERCICE 12-7.24 Montrer que pour tout $a > 0$ on a : $\sum_{k=0}^{+\infty} \frac{(-1)^k}{ak+1} = \int_0^1 \frac{dx}{1+x^a}$.

Chapitre 13

Formes quadratiques sur un espace réel

Nous nous limiterons dans ce chapitre à l'étude des formes quadratiques sur un espace vectoriel réel. De très nombreux résultats subsistent lorsqu'on travaille sur un corps quelconque (de caractéristique différente de 2). Seuls les résultats relatifs au signe d'une forme quadratique sont spécifiques à $\mathbb{R}$.

13-1 Formes bilinéaires et formes quadratiques

13-1.1 Définitions

13-1.1.1 Formes bilinéaires symétriques et antisymétriques

Conformément à ce qui a déjà été vu dans le chapitre traitant des applications multilinéaires et des déterminants :

DÉFINITION 13-1.1 *Si $\mathbb{E}$ est un $\mathbb{R}$-ev, une forme bilinéaire sur l'espace $\mathbf{E}$ est une application*

$$\varphi : \mathbb{E} \times \mathbb{E} \to \mathbb{R} \quad (x, y) \mapsto \varphi(x, y)$$

linéaire par rapport à chacun de ses arguments :

$$\forall x \in \mathbf{E} \quad \varphi(x, \bullet) \ \text{et}\ \varphi(\bullet, x) \in \mathbb{E}^* \ \text{(espace dual de } \mathbb{E}\text{)}$$

On sait que l'espace $\mathcal{L}_2(\mathbb{E},\mathbb{R})$ des formes bilinéaires sur $\mathbb{E}$ est un espace vectoriel (c'est clairement un sous-espace vectoriel de $\mathbb{R}^{\mathbb{E}\times\mathbb{E}}$).

DÉFINITION 13-1.2 *Une forme bilinéaire φ sur $\mathbb{E}$ est symétrique si et seulement si*

$$\forall (x,y)\in\mathbb{E}^2 \quad \varphi(x,y)=\varphi(y,x)$$

Elle est antisymétrique si et seulement si

$$\forall (x,y)\in\mathbb{E}^2 \quad \varphi(x,y)=-\varphi(y,x)$$

On note $\mathcal{L}_{2,s}(\mathbb{E},\mathbb{R})$ (resp. $\mathcal{L}_{2,a}(\mathbb{E},\mathbb{R})$) l'ensemble des formes bilinéaires symétriques (resp. antisymétriques) sur l'espace $\mathbb{E}$.

PROPOSITION 13-1.3 $\mathcal{L}_{2,s}(\mathbb{E},\mathbb{R})$ **et** $\mathcal{L}_{2,a}(\mathbb{E},\mathbb{R})$ **sont deux sous-espaces supplémentaires de** $\mathcal{L}_2(\mathbb{E},\mathbb{R})$. **Toute forme** $\varphi\in\mathcal{L}_2(\mathbb{E},\mathbb{R})$ **se décompose de manière unique sous la forme**

$$\varphi=\varphi_s+\varphi_a \quad \textbf{avec } \varphi_s\in\mathcal{L}_{2,s}(\mathbb{E},\mathbb{R}) \textbf{ et } \varphi_a\in\mathcal{L}_{2,a}(\mathbb{E},\mathbb{R})$$

et on a, pour tout couple (x,y) **de vecteurs de** $\mathbb{E}$,

$$\varphi_s(x,y)=\frac{1}{2}[\varphi(x,y)+\varphi(y,x)] \quad \textbf{et} \quad \varphi_a(x,y)=\frac{1}{2}[\varphi(x,y)-\varphi(y,x)]$$

Démonstration : Exercice. Le résultat subsiste si on remplace le corps des réels par un corps de caractéristique $\neq 2$. ■

Comme nous l'avons montré dans le chapitre sur les déterminants :

PROPOSITION 13-1.4 Une forme bilinéaire φ **sur** $\mathbb{E}$ **est antisymétrique si et seulement si**

$$\forall x\in\mathbb{E} \quad \varphi(x,x)=0$$

EXEMPLE 13-1.5 Si l_1 et l_2 sont deux formes linéaires sur l'espace $\mathbb{E}$, montrer que l'application

$$\varphi:\mathbb{E}\times\mathbb{E}\to\mathbb{R} \quad (x,y)\mapsto\varphi(x,y)=l_1(x)\,l_2(y)$$

est une forme bilinéaire sur $\mathbb{E}$. A quelle condition est-elle symétrique ? antisymétrique ? En particulier, si l'espace $\mathbb{E}$ est de dimension finie, en travaillant avec les formes coordonnées dans une base $\mathcal{B}=\{e_1,\ldots,e_n\}$, on voit que toute application

$$\left(x=\sum_{i=1}^n x_ie_i,\, y=\sum_{j=1}^n y_je_j\right)\mapsto\sum_{i,j}a_{i,j}x_iy_j$$

(où $A=(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}\in\mathcal{M}_n(\mathbb{R})$) est combinaison linéaire de formes bilinéaires et est donc également dans $\mathcal{L}_2(\mathbb{E},\mathbb{R})$. Nous verrons ultérieurement qu'en dimension finie toute forme bilinéaire est de ce type.

EXEMPLE 13-1.6 Sur l'espace $\mathcal{C}_m^0([a,b],\mathbb{R})$ des fonctions réelles continues par morceaux sur un segment $[a,b]$, l'application

$$(f,g)\mapsto\int_a^b f(t)\,g(t)\,\omega(t)\,dt$$

où $\omega\in\mathcal{C}_m^0([a,b],\mathbb{R})$ est fixée, est clairement une forme bilinéaire symétrique.

13-1.1.2 Forme quadratique associée à une forme bilinéaire symétrique

DÉFINITION 13-1.7 *Une application $\Phi : \mathbb{E} \to \mathbb{R}$ est une forme quadratique si et seulement s'il existe une forme bilinéaire symétrique sur $\mathbb{E}$ telle que*

$$\forall x \in \mathbf{E} \quad \Phi(x) = \varphi(x, x)$$

On dit alors que Φ est la forme quadratique associée à la forme bilinéaire φ.

L'espace $Q(\mathbb{E})$ des formes quadratiques sur $\mathbb{E}$ est clairement un espace vectoriel (sous-espace vectoriel de $\mathbf{E}^{\mathbf{R}}$) et l'application

$$\mathcal{L}_{2,s}(\mathbb{E}, \mathbb{R}) \to Q(\mathbb{E}) \quad \varphi \mapsto \Phi \text{ définie sur } \mathbb{E} \text{ par } \Phi(x) = \varphi(x, x)$$

est évidemment linéaire. C'est en fait un isomorphisme d'espace vectoriels :

THÉORÈME 13-1.8 Si $\Phi \in Q(\mathbb{E})$, il existe une unique forme bilinéaire symétrique à laquelle Φ soit associée. Cette forme bilinéaire symétrique φ est appelée forme polaire de la forme quadratique Φ. Elle est reliée à celle-ci par la formule (dite de polarisation)

$$\forall (x, y) \in \mathbb{E}^2 \quad \varphi(x, y) = \frac{1}{2}[\Phi(x + y) - \Phi(x) - \Phi(y)]$$

La correspondance $\varphi \mapsto \Phi$ est donc un isomorphisme de $\mathcal{L}_{2,s}(\mathbb{E}, \mathbb{R})$ dans $Q(\mathbb{E})$.

Démonstration : Si Φ est associée à une forme bilinéaire symétrique φ, on a

$$\forall (x, y) \in \mathbb{E}^2 \quad \Phi(x + y) = \varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y)$$

en développant par bilinéarité et en tenant compte de la symétrie de φ. On a donc bien l'expression annoncée de $\varphi(x, y)$ en fonction de $\Phi(x + y)$, $\Phi(x)$ et $\Phi(y)$. ■

En développant $\varphi(\lambda x + y, \lambda x + y)$ où x et y sont deux vecteurs de $\mathbb{E}$ et λ un scalaire, on obtient l'identité remarquable :

$$\Phi(\lambda x + y) = \lambda^2 \Phi(x) + 2\lambda\varphi(x, y) + \Phi(y)$$

On en déduit aisément une seconde égalité permettant de reconstruire la forme polaire à partir de la forme quadratique :

$$\forall (x, y) \in \mathbf{E}^2 \quad \varphi(x, y) = \frac{1}{4}[\Phi(x + y) - \Phi(x - y)]$$

Remarquons également que, par construction, une forme quadratique $\Phi \in Q(\mathbb{E})$ est une application homogène de degré 2 de $\mathbb{E}$ dans $\mathbb{R}$:

$$\forall x \in \mathbf{E} \quad \forall \lambda \in \mathbf{R} \quad \Phi(\lambda x) = \lambda^2 \Phi(x)$$

EXERCICE 13-1.9 Montrer qu'une application Q de $\mathbb{E}$ dans $\mathbb{R}$ est une forme quadratique sur l'espace $\mathbb{E}$ si et seulement si elle est homogène de degré 2 et si l'application

$$(x,y) \mapsto Q(x+y) - Q(x) - Q(y)$$

est bilinéaire (comme cette application est symétrique, on a simplement à vérifier la linéarité par rapport à un des arguments).

REMARQUE 13-1.10 Si $\psi \in \mathcal{L}_2(\mathbb{E},\mathbb{R})$ est une forme bilinéaire (non nécessairement symétrique) l'application $\Psi : \mathbb{E} \to \mathbb{R}$ définie par $\Psi(x) = \psi(x,x)$ est aussi une forme quadratique. Pour s'en convaincre, il suffit de décomposer $\psi = \psi_s + \psi_a$ en partie symétrique et antisymétrique et de remarquer que, pour tout x de $\mathbb{E}$, on a $\psi(x,x) = \psi_s(x,x)$, à cause de l'antisymétrie de ψ_a. La correspondance $\psi \mapsto \Psi$ est toujours linéaire de $\mathcal{L}_2(\mathbb{E},\mathbb{R})$ vers $\mathcal{Q}(\mathbb{E})$, mais le noyau de ce morphisme est $\mathcal{L}_{2,a}(\mathbb{E},\mathbb{R})$.

13-1.1.3 Restriction à un sous-espace

Soit $\Phi \in \mathcal{Q}(\mathbb{E})$ une forme quadratique sur $\mathbb{E}$. Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}$, la restriction de Φ à $\mathbb{F}$ est une forme quadratique sur $\mathbb{F}$: il suffit en effet de remarquer que la restriction de la forme polaire (à $\mathbb{F} \times \mathbb{F}$)

$$\psi = \varphi_{\mathbb{F}\times\mathbb{F}} : \mathbb{F} \times \mathbb{F} \to \mathbb{R}$$

est une forme bilinéaire symétrique sur $\mathbb{F}$, vérifiant évidemment $\Phi|_{\mathbb{F}}(x) = \psi(x,x)$ pour tout $x \in \mathbb{F}$.

13-1.2 Forme quadratique positive, définie positive

Dans cette section, le fait que le corps de base soit $\mathbb{R}$ est évidemment fondamental.

13-1.2.1 Définitions

DÉFINITION 13-1.11 *Une forme quadratique $\Phi \in \mathcal{Q}(\mathbb{E})$ est positive sur $\mathbb{E}$ si et seulement si*

$$\forall x \in \mathbb{E} \quad \Phi(x) \geqslant 0$$

Elle est négative sur $\mathbb{E}$ si $-\Phi$ est positive. On dit enfin que Φ est définie[1] *sur $\mathbb{E}$ si et seulement si*

$$\forall x \neq 0_{\mathbb{E}} \quad \Phi(x) \neq 0$$

PROPOSITION 13-1.12 Si Φ est une forme quadratique définie sur $\mathbb{E}$, elle est définie positive ou définie négative[2]**.**

---

[1] Terminologie usuelle, mais un peu malheureuse.

[2] Sur un corps ordonné quelconque, ce résultat n'est plus vrai : sur $\mathbb{E} = \mathbb{Q}^2$ considéré comme $\mathbb{Q}$-ev , l'application

$$\mathbb{E} \to \mathbb{Q} \quad (x,y) \mapsto x^2 - 2y^2$$

est une forme quadratique définie, qui ne garde pas un signe constant.

Démonstration : Supposons qu'il existe deux vecteurs x et $y \in \mathbf{E}$ tels que

$$\Phi(x) > 0 \quad \text{et} \quad \Phi(y) < 0$$

Pour $t \in \mathbb{R}$, on considère

$$\Phi(t\,x + y) = t^2\Phi(x) + 2t\,\varphi(x,y) + \Phi(y)$$

expression polynomiale de degré 2, positive au voisinage de $+\infty$ et négative en 0. On en déduit l'existence d'un réel $t_0 > 0$ avec

$$\Phi(t_0 x + y) = 0$$

soit $t_0 x + y = 0_{\mathbf{E}}$ puisque Φ est supposée définie. Mais on a alors $\Phi(y) = \Phi(-t_0 x) = t_0^2\Phi(x) > 0$, ce qui contredit $\Phi(y) < 0$. Tous les vecteurs de $\mathbf{E}$ non nuls ont donc des images par Φ de même signe. ■

Une forme quadratique définie sur $\mathbb{E}$ pourra donc toujours être supposée positive (après multiplication éventuelle par -1). Nous noterons $\mathcal{Q}^+(\mathbb{E})$ l'ensemble des formes quadratiques positives sur $\mathbb{E}$ et $\mathcal{Q}^{*+}(\mathbb{E})$ celui des formes quadratiques définies positives sur $\mathbb{E}$. Il s'agit de deux cônes de $\mathcal{Q}(\mathbb{E})$ (c'est-à-dire d'ensembles stables par les homothéties de rapports strictement positifs) qui sont clairement convexes.

REMARQUE 13-1.13 Par abus de langage, nous dirons parfois qu'une forme bilinéaire symétrique φ sur $\mathbb{E}$ est définie positive lorsque la forme quadratique associée l'est. Il est cependant évident qu'une forme bilinéaire symétrique non nulle $\varphi : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ est surjective, et $\varphi(\mathbb{E} \times \mathbb{E})$ ne peut pas être inclus dans $\mathbb{R}^+$.

13-1.2.2 Inégalité de Cauchy-Schwarz

PROPOSITION 13-1.14 Si Φ est une forme quadratique positive sur $\mathbb{R}$ et φ est sa forme polaire, on a

$$\forall x, y \in \mathbb{E} \quad [\varphi(x,y)]^2 \leqslant \Phi(x).\Phi(y)$$

Il y a égalité si et seulement s'il existe une combinaison linéaire z de x et y à coefficients non tous nuls telle que $\Phi(z) = 0$. Lorsque Φ est définie positive, il y a égalité si et seulement si $\{x, y\}$ est liée.

Démonstration : Pour $t \in \mathbb{R}$, on pose, x et y étant deux vecteurs de $\mathbb{E}$,

$$P(t) = \Phi(t\,x + y) = t^2\Phi(x) + 2t\varphi(x,y) + \Phi(y)$$

fonction polynomiale de degré inférieur à 2, ne prenant par hypothèse que des valeurs positives. Si $\Phi(x) = 0$, ce polynôme ne peut être que constant, donc $\varphi(x,y) = 0$, et l'inégalité à prouver est alors une égalité. Si $\Phi(x) > 0$, le polynôme est de degré 2 et ne garde un signe constant que si son discriminant est négatif ou nul, ce qui donne bien

$$\varphi^2(x,y) \leqslant \Phi(x).\Phi(y)$$

S'il y a égalité dans l'inégalité de Schwarz le polynôme P est constant donc $\Phi(x) = 0$ ou P est de degré 2 à discriminant nul, et donc il existe $t_0 \in \mathbb{R}$ tel que $\Phi(z) = 0$, pour $z = t_0 x + y$. Réciproquement, lorsque $\Phi(x) \neq 0$, l'existence d'un tel t_0 montre que le polynôme P possède une racine. Comme P est positif, cette racine est nécessairement double, et le discriminant de P est nul, ce qui donne l'inégalité de Schwarz. ■

REMARQUE 13-1.15 L'inégalité de Schwarz peut aussi s'écrire

$$\forall x, y \in \mathbb{E} \quad |\varphi(x,y)| \leqslant \sqrt{\Phi(x)}\sqrt{\Phi(y)}$$

Il ne faut évidemment pas mélanger les deux écritures : dans la première les deux membres sont des fonctions homogènes de degré 2 de x. Ce sont des fonctions (positivement) homogènes de degré 1 pour la seconde.

COROLLAIRE 13-1.16 Si $\Phi \in \mathcal{Q}^+(\mathbb{E})$, l'application

$$N_\Phi : \mathbb{E} \to \mathbb{R}^+ \quad x \mapsto \sqrt{\Phi(x)}$$

est une semi-norme sur l'espace $\mathbb{E}$. C'est une norme lorsque Φ est définie positive.

Démonstration : On a évidemment

$$\forall x \in \mathbb{E} \quad \forall \lambda \in \mathbb{R} \quad N_\Phi(\lambda x) = |\lambda| . N_\Phi(x)$$

L'équivalence $N_\Phi(x) = 0 \Leftrightarrow x = 0_{\mathbb{E}}$ n'est évidemment vérifiée que lorsque l'on a $\Phi \in \mathcal{Q}^{*+}(\mathbb{E})$. Enfin, l'inégalité triangulaire

$$\forall (x,y) \in \mathbb{E}^2 \quad \sqrt{\Phi(x+y)} \leqslant \sqrt{\Phi(x)} + \sqrt{\Phi(y)}$$

(appelée dans ce cas inégalité de Minkowski) est conséquence de l'inégalité de Schwarz : s'agissant d'une inégalité entre réels positifs, elle sera vérifiée si et seulement si les carrés des deux membres sont dans le même ordre, soit

$$\Phi(x+y) \leqslant \Phi(x) + \Phi(y) + 2\sqrt{\Phi(x)}\sqrt{\Phi(y)}$$

En développant par bilinéarité le premier membre, cette inégalité est équivalente à

$$\varphi(x,y) \leqslant \sqrt{\Phi(x)}\sqrt{\Phi(y)}$$

ce qui est bien vérifié par l'inégalité de Cauchy-Schwarz. ■

COROLLAIRE 13-1.17 L'inégalité de Minkowski

$$\sqrt{\Phi(x+y)} \leqslant \sqrt{\Phi(x)} + \sqrt{\Phi(y)}$$

est une égalité si et seulement s'il y a égalité dans l'inégalité de Schwarz et si $\varphi(x,y) \geqslant 0$. Lorsque $\Phi \in \mathcal{Q}^{*+}(\mathbb{E})$, cela signifie exactement que l'un des deux vecteurs x ou y est nul ou que x et y sont $\mathbb{R}^{*+}$ proportionnels (donc colinéaires et de même sens).

L'inégalité de Cauchy-Schwarz a une conséquence importante, sur laquelle nous reviendrons dans le cas particulier de la dimension finie :

COROLLAIRE 13-1.18 Si Φ est une forme quadratique positive sur $\mathbb{E}$ et de forme polaire φ, pour $x \in \mathbf{E}$ on a

$$\Phi(x) = 0 \Leftrightarrow \varphi(x, \bullet) = 0_{\mathbb{E}^*}$$

(forme linéaire nulle sur l'espace $\mathbb{E}$).

> *Démonstration* : l'implication $\varphi(x, \bullet) = 0_{\mathbb{E}^*} \Rightarrow \Phi(x) = 0$ est toujours vraie, sans considération de signe pour Φ, tout simplement parce que $\Phi(x) = \varphi(x, x)$. La réciproque est conséquence immédiate de l'inégalité
>
> $$\forall y \in \mathbb{E} \quad |\varphi(x, y)| \leqslant \sqrt{\Phi(x)}\sqrt{\Phi(y)} \quad \blacksquare$$

Attention ! Ce corollaire n'est plus vrai pour une forme quadratique dont le signe n'est pas constant : on vérifie par exemple aisément que

$$\Phi : \mathbb{R}^2 \to \mathbb{R} \quad x = (x_1, x_2) \mapsto x_1^2 - x_2^2$$

est une forme quadratique[3] sur $\mathbf{E} = \mathbb{R}^2$, et que le vecteur $x = (1, 1)$ annule Φ, alors que

$$\varphi(x, \bullet) : \mathbb{R}^2 \to \mathbb{R} \quad (y_1, y_2) \mapsto y_1 - y_2$$

n'est pas la forme nulle.

13-2 Formes quadratiques en dimension finie

Dans cette section, l'espace $\mathbb{E}_n$ est supposé de dimension n.

13-2.1 Matrice d'une forme bilinéaire dans une base

13-2.1.1 Définition

Supposons que $\varphi \in \mathcal{L}_2(\mathbb{E}_n, \mathbb{R})$ soit une forme bilinéaire sur $\mathbb{E}_n$. Si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base de $\mathbb{E}_n$, on pourra développer $\varphi(x, y)$ par bilinéarité si l'on connaît les coordonnées de x et y dans $\mathcal{B}$ et les images des couples de vecteurs de base par φ : pour

$$x = \sum_{i=1}^{n} x_i e_i \quad \text{et} \quad y = \sum_{j=1}^{n} y_j e_j$$

[3] Associée à la forme bilinéaire symétrique φ définie par

$$\varphi((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2$$

on aura

$$\varphi(x,y)=\sum_{i=1}^{n}x_i\varphi(e_i,y)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j\varphi(e_i,e_j)$$

On voit ainsi que φ est entièrement déterminée dès que l'on connaît les n^2 images par φ des couples de vecteurs de base. Ceci amène à la définition

DÉFINITION 13-2.1 *Si $\varphi\in\mathcal{L}_2(\mathbb{E}_n,\mathbb{R})$ et $\mathcal{B}=\{e_1,\dots,e_n\}$ est une base de $\mathbb{E}_n$, on appelle matrice de φ dans la base $\mathcal{B}$ la matrice carrée*

$$M=\mathcal{M}(\varphi,\mathcal{B})=(\varphi(e_i,e_j))_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}\in\mathcal{M}_n(\mathbb{R})$$

Le calcul précédemment effectué peut alors s'écrire matriciellement : si X et Y sont les vecteurs colonnes représentant respectivement x et y dans $\mathcal{B}$, le scalaire $\sum_{i=1}^{n}x_i\left(\sum_{j=1}^{n}\varphi(e_i,e_j)\,y_j\right)$ peut s'interpréter comme le produit de la ligne ${}^tX$ par la colonne MY et par conséquent

$$\varphi(x,y)={}^tXMY$$

13-2.1.2 Dimension des espaces $\mathcal{L}_2(\mathbb{E}_n,\mathbb{R})$ et $\mathcal{L}_{2,s}(\mathbb{E}_n,\mathbb{R})$

Nous avons vu que la matrice d'une forme bilinéaire dans une base caractérisait entièrement cette forme bilinéaire. Plus précisément :

PROPOSITION 13-2.2 Si $\mathcal{B}$ est une base fixée de $\mathbb{E}_n$, l'application

$$\mathcal{L}_2(\mathbb{E}_n,\mathbb{R})\to\mathcal{M}_n(\mathbb{R})\quad\varphi\mapsto\mathcal{M}(\varphi,\mathcal{B})$$

est un isomorphisme d'espace vectoriel.

> *Démonstration* : Cette application est clairement linéaire. On sait qu'elle est injective. Si $A=(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}\in\mathcal{M}_n(\mathbb{R})$ est donnée, l'expression
>
> $$\varphi(x,y)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{i,j}x_iy_j$$
>
> (avec bien sûr $x=\sum_{i=1}^{n}x_ie_i$ et $y=\sum_{j=1}^{n}y_je_j$) définit une forme bilinéaire sur $\mathbb{E}_n$ (voir exemple **13-1**.5) vérifiant $\varphi(e_i,e_j)=a_{i,j}$, donc $\mathcal{M}(\varphi,\mathcal{B})=A$. L'application considérée est donc aussi surjective. ■

COROLLAIRE 13-2.3 Avec les notations qui précèdent, φ est symétrique si et seulement si $M=\mathcal{M}(\varphi,\mathcal{B})$ l'est, c'est-à-dire

$${}^tM=M$$

et l'application $\mathcal{L}_{2,s}(\mathbb{E}_n,\mathbb{R}) \to \mathcal{S}_n(\mathbb{R})$ qui à une forme bilinéaire symétrique associe sa matrice dans $\mathcal{B}$ est un isomorphisme d'espaces vectoriels ($\mathcal{S}_n(\mathbb{R})$ désignant l'espace des matrices carrées réelles symétriques d'ordre n). De même φ est antisymétrique ssi

$$^tM = -M$$

et $\mathcal{L}_{2,a}(\mathbb{E}_n,\mathbb{R})$ est isomorphe à $\mathcal{A}_n(\mathbb{R})$ espace des matrices carrées antisymétriques.

Démonstration : Si φ est symétrique, on a évidemment

$$\forall i,j \quad \varphi(e_i,e_j) = \varphi(e_j,e_i)$$

et donc $^tM = M$. Si réciproquement M est symétrique, on a en calculant dans la base $\mathcal{B}$

$$\forall x,y \in \mathbb{E} \quad \varphi(x,y) = {}^tXMY = {}^t\left({}^tXMY\right) = {}^tY{}^tMX = {}^tYMX = \varphi(y,x)$$

Le cas de l'antisymétrie se traite de la même manière. Remarquons que, matriciellement, la décomposition d'une forme bilinéaire en somme d'une partie symétrique et d'une partie antisymétrique est simplement la décomposition

$$M = \frac{1}{2}\left(M + {}^tM\right) + \frac{1}{2}\left(M - {}^tM\right)$$

de M dans la somme directe $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$. ■

COROLLAIRE 13-2.4 Si $\mathbb{E}_n$ est un espace vectoriel de dimension n

$$\dim \mathcal{L}_{2,s}(\mathbb{E}_n,\mathbb{R}) = \frac{n(n+1)}{2} \quad \textbf{et} \quad \dim \mathcal{L}_{2,a}(\mathbb{E}_n,\mathbb{R}) = \frac{n(n-1)}{2}$$

(et donc évidemment $\dim \mathcal{L}_2(\mathbb{E}_n,\mathbb{R}) = n^2$).

REMARQUE 13-2.5 Attention aux notations : ne pas confondre la notion de matrice d'un endomorphisme dans une base avec celle de matrice d'une forme bilinéaire. La matrice d'une forme bilinéaire se remplit coefficient par coefficient ; si on change par exemple un des vecteurs de la base, on ne modifie qu'une ligne et une colonne de la matrice. C'est totalement différent de ce qui se passe avec la matrice d'un endomorphisme, qui est en général complètement transformée dès qu'on change un vecteur de base. Nous verrons dans la section suivante que les formules de changement de base ne sont pas les mêmes.

REMARQUE 13-2.6 Dans le même ordre d'idée, si $\mathbb{E}_n = \mathbb{F}_p \oplus \mathbb{G}_{n-p}$ et si on choisit $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ une base adaptée à cette décomposition on aura une décomposition par blocs de la matrice d'une forme bilinéaire

$$M = \mathcal{M}(\varphi,\mathcal{B}) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

avec immédiatement

$$A = \mathcal{M}\left(\varphi|_{\mathbb{F}_p\times\mathbb{F}_p}, \mathcal{B}_1\right) \in \mathcal{M}_p(\mathbb{R}) \quad \text{et} \quad D = \mathcal{M}\left(\varphi|_{\mathbb{G}_{n-p}\times\mathbb{G}_{n-p}}, \mathcal{B}_2\right)$$

On voit en particulier que la donnée des restrictions de φ aux deux sous-espaces supplémentaires $\mathbb{F}_p$ et $\mathbb{G}_{n-p}$ ne permet pas de reconstruire φ en totalité : en travaillant matriciellement, on n'a aucune information sur les matrices B et C.

REMARQUE 13-2.7 Si φ est une forme bilinéaire sur $\mathbb{E}_n$ et Φ est la forme quadratique associée, la matrice de φ dans $\mathcal{B}$ sera aussi appelée matrice de Φ dans cette base (il n'y a aucune ambiguité, puisque Φ caractérise φ) :

$$\mathcal{M}(\Phi, \mathcal{B}) = (\varphi(e_i, e_j))_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} \in \mathcal{S}_n(\mathbb{R})$$

Ici encore, il est une erreur à ne pas commettre : les coefficients diagonaux de cette matrice sont les $(\Phi(e_i))_{1\leqslant i\leqslant n}$. Ils donnent certes une information sur Φ. Mais ces n coefficients sont insuffisants pour déterminer Φ : **une forme quadratique n'est pas caractérisée par les images des vecteurs d'une base de l'espace !**

13-2.1.3 Effet d'un changement de base : matrices congruentes

Soit $\varphi \in \mathcal{L}_2(\mathbb{E}_n, \mathbb{R})$ et M sa matrice dans une base $\mathcal{B}$ de $\mathbb{E}_n$. Si $\mathcal{B}'$ est une autre base et $P \in \mathcal{GL}_n(\mathbb{R})$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, on détermine à l'aide de M et P la matrice N de φ dans $\mathcal{B}'$: si x et y sont deux vecteurs de $\mathbb{E}_n$ représentés par les colonnes X et Y dans $\mathcal{B}$ et les colonnes X' et Y' dans $\mathcal{B}'$, on a d'après les formules de changement de bases

$$X = PX' \text{ et } Y = PY'$$

ce qui donne

$$\varphi(x,y) = {}^tXMY = {}^t(PX')M(PY') = {}^tX'\left({}^tPMP\right)Y'$$

Cette égalité étant vérifiée quels que soient les vecteurs X' et Y' on a, d'après l'unicité de l'écriture matricielle d'une forme bilinéaire dans une base donnée (proposition **13-2**.2),

$$N = {}^tPMP$$

DÉFINITION 13-2.8 *Si M et N sont deux matrices de $\mathcal{M}_n(\mathbb{R})$, on dit que N est congruente à M si et seulement s'il existe une matrice $P \in \mathcal{GL}_n(\mathbb{R})$ telle que $N = {}^tPMP$.*

Avec le calcul précédent, en interprétant par exemple M comme matrice d'une forme bilinéaire dans la base canonique de $\mathbb{R}^n$ et P comme matrice de passage de cette base canonique à une autre base de $\mathbb{R}^n$, on obtient immédiatement

PROPOSITION 13-2.9 Si M et N sont dans $\mathcal{M}_n(\mathbb{R})$, N est congruente à M si et seulement si M et N représentent la même forme bilinéaire dans deux bases de $\mathbb{R}^n$ (ou de tout espace réel de dimension n). La relation de congruence est une relation d'équivalence sur $\mathcal{M}_n(\mathbb{R})$.

(cette dernière affirmation peut aussi se vérifier directement). Il est clair que si M est symétrique, toute matrice congruente à M l'est aussi : on peut utiliser la proposition précédente, ou utiliser simplement

$$^t\left({}^tPMP\right) = {}^tP\,{}^tMP$$

On a une propriété analogue dans le cas de matrices antisymétriques.

Attention ! Contrairement à ce qui se passe pour les matrices semblables, les déterminants de deux matrices congruentes ne sont pas égaux en général :

$$\forall M, N \in \mathcal{M}_n(\mathbb{R}) \quad \forall P \in \mathcal{GL}_n(\mathbb{R})$$
$$N = {}^tPMP \Rightarrow \det(N) = [\det(P)]^2 . \det(M)$$

On s'aperçoit simplement que la nullité de l'un des déterminants équivaut à la nullité de l'autre, et que, lorsqu'ils sont non nuls, ces deux déterminants ont même signe (puisque $[\det(P)]^2 > 0$).

DÉFINITION 13-2.10 *Si Φ est une forme quadratique sur $\mathbb{E}_n$, le discriminant de Φ dans une base de $\mathbb{E}_n$ est le déterminant de la matrice de Φ dans cette base. Le signe de ce discriminant ne dépend pas de la base choisie.*

Il est également une propriété conservée par congruence, sur laquelle nous reviendrons dans le cas des matrices symétriques :

PROPOSITION 13-2.11 Deux matrices congruentes ont même rang.

13-2.2 Expression d'une forme quadratique dans une base en dimension finie

Revenons au calcul effectué à la section **13-2.1.1** dans le cas d'une forme bilinéaire symétrique et de la forme quadratique qui lui est associée : si $M = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} \in \mathcal{S}_n(\mathbb{R})$ est la matrice de $\varphi \in \mathcal{L}_{2,s}(\mathbb{E}_n, \mathbb{R})$ dans une base $\mathcal{B}$, et si X et Y sont les colonnes représentatives de x et $y \in \mathbb{E}_n$ dans cette base, on a

$$\varphi(x, y) = {}^tXMY = \sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} a_{i,j} x_i y_j$$

ce qui donne l'expression de la forme quadratique

$$\forall x \in \mathbb{E}_n \quad \Phi(x) = {}^tXMX = \sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} a_{i,j} x_i x_j$$

qu'on peut écrire aussi, en tenant compte de la symétrie de la matrice M

$$\Phi(x) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{\substack{(i,j)\\ 1\leqslant i<j\leqslant n}} a_{ij} x_i x_j$$

L'usage veut que l'on désigne un terme comme "$a_{ii}x_i^2$" sous l'appellation "terme carré", alors que, pour $i \neq j$, $2a_{ij}x_ix_j$ est un "terme rectangle". Dans la base $\mathcal{B}$, $\Phi(x)$ s'exprime donc comme fonction polynomiale homogène de degré 2 des coordonnées du vecteur x. Réciproquement, toute expression de ce type définit une forme quadratique sur $\mathbb{E}_n$: si

$$\forall x \in \mathbb{E}_n \quad \Phi(x) = \sum_{k=1}^{n} \alpha_k x_k^2 + \sum_{\substack{(i,j)\\ 1\leqslant i<j\leqslant n}} \beta_{i,j} x_i x_j$$

est un polynôme homogène de degré 2, on peut écrire, avec les mêmes notations que précédemment

$$\forall x \in \mathbb{E}_n \quad \Phi(x) = {}^tXAX$$

où la matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} \in \mathcal{S}_n(\mathbb{R})$ est définie par

$$\forall i \in \{1,\dots,n\} \quad a_{ii} = \alpha_i \quad \text{et} \quad \forall i,j \in \{1,\dots,n\} \text{ avec } i<j \quad a_{ij} = a_{ji} = \frac{1}{2}\beta_{i,j}$$

et il est clair que $\varphi(x,y) = {}^tXAY$ définit une forme bilinéaire symétrique sur $\mathbb{E}_n$, et que Φ est la forme quadratique qui lui est associée.

Attention au facteur $\frac{1}{2}$: on dit que l'on dédouble les termes rectangles pour obtenir la matrice de la forme quadratique. L'expression de la forme polaire dans la base $\mathcal{B}$ peut s'obtenir selon le même principe sans qu'il soit nécessaire de présenter les calculs matriciellement[4] :

$$\forall x,y \in \mathbb{E}_n \quad \varphi(x,y) = \sum_{i=1}^{n} \alpha_i x_i y_i + \sum_{\substack{(i,j)\\ 1\leqslant i<j\leqslant n}} \frac{1}{2}\beta_{ij}(x_iy_j + x_jy_i)$$

PROPOSITION 13-2.12 Une application $Q : \mathbb{E}_n \to \mathbb{R}$ est une forme quadratique si et seulement si Q s'exprime dans une base quelconque de $\mathbb{E}_n$ comme fonction polynomiale homogène de degré 2 des coordonnées dans cette base.

Avec le calcul qu'on vient de faire, on a aussi

COROLLAIRE 13-2.13 Si M et N sont deux matrices de $\mathcal{S}_n(\mathbb{R})$

$$M = N \Leftrightarrow \forall X \in \mathbb{R}^n \quad {}^tXMX = {}^tXNX$$

Lorsque M et N sont deux matrices de $\mathcal{M}_n(\mathbb{R})$, la proposition 13-1.4 montre d'ailleurs que

$$\forall X \in \mathbb{R}^n \quad {}^tXMX = {}^tXNX \Leftrightarrow M - N \in \mathcal{A}_n(\mathbb{R})$$

[4] Retenir que le terme carré x_i^2 se "polarise" en x_iy_i, alors qu'un terme rectangle x_ix_j $(i \neq j)$ se "polarise" en $\frac{1}{2}(x_iy_j + x_jy_i)$.

13-2.3 Interprétation de la matrice d'une forme quadratique

13-2.3.1 Morphisme $\mathbb{E}_n \to \mathbb{E}_n^*$ associé à une forme bilinéaire symétrique

Si $\varphi \in \mathcal{L}_{2,s}(\mathbb{E}_n, \mathbb{R})$, pour tout $x \in \mathbb{E}_n$, l'application

$$\varphi(x, \bullet) : \mathbb{E}_n \to \mathbb{R} \quad y \mapsto \varphi(x, y)$$

est une forme linéaire sur $\mathbb{E}_n$. La linéarité de φ par rapport à son premier argument montre que cette forme dépend linéairement de x.

DÉFINITION 13-2.14 *Si $\varphi \in \mathcal{L}_{2,s}(\mathbb{E}_n, \mathbb{R})$, l'application $\widetilde{\varphi}$ de l'espace $\mathbb{E}_n$ dans son dual $\mathbb{E}_n^*$ définie par*

$$\mathbb{E}_n \ni x \mapsto \widetilde{\varphi}(x) = \varphi(x, \bullet)$$

est un morphisme d'espaces vectoriels, appelé morphisme de $\mathbb{E}_n$ vers $\mathbb{E}_n^*$ canoniquement associé à φ.

Nous pouvons à présent interpréter la matrice de φ dans une base comme matrice d'un morphisme d'espace vectoriel. Si $\mathcal{B}$ est une base de $\mathbb{E}_n$, on note $\mathcal{B}^*$ sa base duale (c'est-à-dire la base de $\mathbb{E}_n^*$ composée des formes coordonnées dans $\mathcal{B}$).

PROPOSITION 13-2.15 Si $\varphi \in \mathcal{L}_{2,s}(\mathbb{E}_n, \mathbb{R})$ et $\mathcal{B}$ est une base de $\mathbb{E}_n$

$$\mathcal{M}(\varphi, \mathcal{B}) = \mathcal{M}(\widetilde{\varphi}, \mathcal{B}, \mathcal{B}^*)$$

Démonstration : Le coefficient $a_{i,j}$ d'indices (i, j) de la matrice $A = \mathcal{M}(\widetilde{\varphi}, \mathcal{B}, \mathcal{B}^*)$ est, par définition, la $i^{\text{ème}}$ coordonnée dans la base $\mathcal{B}^*$ de $\widetilde{\varphi}(e_j) = \varphi(e_j, \bullet)$, c'est-à-dire, par définition de la base duale

$$\widetilde{\varphi}(e_j)(e_i) = \varphi(e_j, e_i) = \varphi(e_i, e_j)$$

puisque φ est symétrique. ■

EXERCICE 13-2.16 Utiliser ce résultat pour retrouver la formule de changement de base obtenue à la section **13-2.1.3**

En utilisant cette interprétation de la matrice d'une forme quadratique, on comprend mieux pourquoi son rang ne dépend pas de la base choisie pour écrire la matrice :

13-2.3.2 Rang et radical d'une forme quadratique

DÉFINITION 13-2.17 *Soit $\Phi \in \mathcal{Q}(\mathbb{E}_n)$ et $\varphi \in \mathcal{L}_{2,s}(\mathbb{E}_n, \mathbb{R})$ sa forme polaire. On appelle rang de Φ ou rang de φ le rang de l'application $\widetilde{\varphi}$. C'est donc aussi le rang de la matrice de φ dans une base quelconque de $\mathbb{E}_n$:*

$$\operatorname{rang} \Phi = \operatorname{rang} \varphi = \operatorname{rang} \widetilde{\varphi} = \operatorname{rang} \mathcal{M}(\varphi, \mathcal{B})$$

Il est intéressant d'étudier le noyau du morphisme $\widetilde{\varphi}$: c'est la notion de radical d'une forme quadratique.

DÉFINITION 13-2.18 *Si $\Phi \in \mathcal{Q}(\mathbb{E}_n)$ et φ est sa forme polaire, on appelle radical*[5] *de Φ (ou de φ) le noyau du morphisme*[6] $\widetilde{\varphi}$ *:*

$$\mathrm{Rad}\,\Phi = \ker \widetilde{\varphi} = \{x \in \mathbb{E}_n \mid \varphi(x, \bullet) = 0_{\mathbb{E}_n^*}\}$$

Le théorème du rang appliqué à $\widetilde{\varphi}$ donne immédiatement :

$$\dim \mathrm{Rad}\,\Phi = \dim \mathbb{E}_n - \mathrm{rang}\,\Phi$$

Lorsque le rang de Φ est strictement inférieur à la dimension de l'espace, on dit que la forme quadratique Φ est **dégénérée** : cela veut dire qu'en fait, l'étude de Φ peut être ramenée à celle d'une forme quadratique sur un supplémentaire de $\mathrm{Rad}\,\Phi$, c'est-à-dire sur un espace de dimension $< n$. En effet, si $\mathbb{E}_n = \mathbb{S} \oplus \mathrm{Rad}\,\Phi$, pour x et y se décomposant en $x = x_{\mathbb{S}} + x_R$ et $y = y_{\mathbb{S}} + y_H$ dans cette somme directe, on a

$$\varphi(x, y) = \varphi(x_{\mathbb{S}}, y_{\mathbb{S}}) + \varphi(x_{\mathbb{S}}, y_R) + \varphi(x_R, y_{\mathbb{S}}) + \varphi(x_R, y_R) = \varphi(x_{\mathbb{S}}, y_{\mathbb{S}})$$

puisque les formes $\varphi(x_R, \bullet)$ et $\varphi(y_R, \bullet)$ sont nulles. En particulier, pour $x = y$

$$\Phi(x) = \Phi(x_{\mathbb{S}})$$

Matriciellement, si on travaille dans une base $\mathcal{B} = \mathcal{B}_{\mathbb{S}} \cup \mathcal{B}_R$ adaptée à cette décomposition, la matrice de Φ aura la forme

$$M = \mathcal{M}(\Phi, \mathcal{B}) = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$$

où la matrice $A \in \mathcal{M}_r(\mathbb{R})$ (avec $r = \dim \mathbb{S} = \mathrm{rang}\,\Phi$) est la matrice de $\Phi|_{\mathbb{S}}$ dans la base $\mathcal{B}_{\mathbb{S}}$. Comme M doit être de rang r, on a d'ailleurs $A \in \mathcal{GL}_r(\mathbb{R})$. Dans une base bien choisie, certaines coordonnées n'apparaissent pas dans l'expression d'une forme quadratique dégénérée.

13-2.3.3 Forme bilinéaire symétrique et forme quadratique non dégénérée

Conformément à la définition donnée plus haut, on dit qu'une forme quadratique Φ sur $\mathbb{E}_n$ est non dégénérée si elle est de rang n. Cela signifie que sa matrice dans une base quelconque de $\mathbb{E}_n$ est inversible. Ou encore que $\widetilde{\varphi}$ est un isomorphisme de $\mathbb{E}_n$ vers $\mathbb{E}_n^*$:

[5] On dit aussi "noyau de φ" (ou de Φ). Mais cette terminologie est malheureuse, car ce n'est pas l'ensemble

$$\mathcal{C}(\Phi) = \{x \in \mathbb{E}_n \mid \Phi(x) = 0\}$$

On a en effet $\mathrm{Rad}\,\Phi \subset \mathcal{C}(\Phi)$ mais l'inclusion est stricte en général (mais il y a égalité lorsque Φ est positive) : voir proposition **13-1**.18 et le contre-exemple qui suit.

[6] Donc, en travaillant dans une base $\mathcal{B}$, si $M = \mathcal{M}(\Phi, \mathcal{B})$, les vecteurs du radical de Φ sont représentés par les colonnes X vérifiant $MX = 0$.

PROPOSITION 13-2.19 Si Φ est une forme quadratique non dégénérée sur $\mathbb{E}_n$, l'application

$$x \mapsto \varphi(x, \bullet)$$

est un isomorphisme d'espace vectoriel de $\mathbb{E}_n$ sur son dual.

Toute forme linéaire $l \in \mathbb{E}_n^*$ s'écrit donc de manière unique sous la forme $\varphi(x, \bullet)$, avec $x \in \mathbb{E}_n$. Des problèmes faisant intervenir des formes linéaires sur $\mathbb{E}_n$ peuvent donc alors se traiter de manière "interne" (en travaillant avec des vecteurs de l'espace), en identifiant la forme $l = \varphi(x, \bullet)$ avec le vecteur x. Par exemple, un hyperplan $\mathbb{H}$ de $\mathbb{E}_n$ caractérisé par une forme linéaire (ou plus précisément par une droite vectorielle de $\mathbb{E}_n^*$) sera caractérisé par une droite vectorielle de $\mathbb{E}_n$. C'est en particulier ce point de vue que nous adopterons dans le cadre des espaces euclidiens, avec la notion de normale à un hyperplan.

EXEMPLE 13-2.20 Toute forme Φ quadratique définie positive sur $\mathbb{E}_n$ est non dégénérée, puisque dans ce cas

$$\varphi(x, \bullet) = 0_{\mathbb{E}_n^*} \Rightarrow \Phi(x) = \varphi(x, x) = 0 \Rightarrow x = 0_{\mathbb{E}_n}$$

ce qui prouve l'injectivité de $\widetilde{\varphi}$, et donc rang $\Phi = n$. Il n'y a évidemment pas de réciproque. Sur l'espace $\mathbb{E}_2 = \mathbf{R}^2$, la forme quadratique définie sur $X = (x, y)$ par

$$\Phi(X) = x^2 - y^2$$

est non dégénérée[7], mais pas de signe constant.

13-3 Réduction d'une forme quadratique et signature

Nous avons vu dans la section précédente que toute matrice carrée d'ordre n symétrique de rang r était congruente à une matrice de la forme

$$\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \text{ avec } A \in \mathcal{GL}_r(\mathbb{R})$$

Nous allons voir ici que l'on peut aller plus loin, et imposer à A d'être diagonale. Le raisonnement serait valable sur un corps quelconque (de caractéristique $\neq 2$). Si on travaille sur $\mathbf{R}$, nous verrons que, lorsque A est diagonale, les signes de ses coefficients diagonaux ont une interprétation géométrique.

[7] Sa matrice dans la base canonique est $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathcal{GL}_2(\mathbb{R})$. Une forme linéaire quelconque sur l'espace $\mathbb{R}^2$ $l : X = (x, y) \mapsto \alpha x + \beta y$ s'écrit $l = \varphi(X_0, \bullet)$, avec $X_0 = (\alpha, -\beta)$.

13-3.1 Réduction d'une forme quadratique

THÉORÈME 13-3.1 Si Φ est une forme quadratique sur $\mathbb{E}_n$, il existe une base $\mathcal{B}$ de $\mathbb{E}_n$ telle que la matrice de Φ dans $\mathcal{B}$ soit de la forme

$$\mathcal{M}(\Phi, \mathcal{B}) = \text{Diag}(\alpha_1, \dots, \alpha_n)$$

Le nombre r de coefficients diagonaux non nuls est évidemment égal au rang de Φ. En réordonnant les vecteurs de $\mathcal{B}$, on peut donc imposer

$$\mathcal{M}(\Phi, \mathcal{B}) = \text{Diag}(\alpha_1, \dots, \alpha_r, 0, \dots, 0)$$

avec les scalaires $(\alpha_i)_{1 \leqslant i \leqslant r}$ non nuls.

Démonstration : On raisonne par récurrence sur la dimension de l'espace $\mathbb{E}_n$. Pour $n = 1$, il n'y a rien à prouver. On suppose le résultat établi pour toute forme quadratique sur tout espace vectoriel de dimension $n - 1$. Supposons $\Phi \in \mathcal{Q}(\mathbb{E}_n)$. Si Φ est identiquement nulle, la matrice de Φ dans toute base de $\mathbb{E}_n$ est nulle et $r = 0$. Sinon, on peut trouver un vecteur e_1 avec $\Phi(e_1) \neq 0$. Posons $\alpha_1 = \Phi(e_1)$. Si φ est la forme polaire de Φ, la forme linéaire $\psi = \varphi(e_1, \bullet)$ est non nulle, puisque $\psi(e_1) = \alpha_1 \neq 0$. Comme e_1 n'appartient pas à l'hyperplan $\ker\psi$, on a

$$\mathbb{E}_n = \text{vect}(e_1) \oplus \ker\psi$$

et, par l'hypothèse de récurrence appliquée à la forme quadratique $\Phi|_{\ker\psi}$, on peut trouver une base $\mathcal{B}' = \{e_2, \dots, e_n\}$ avec

$$\mathcal{M}(\Phi|_{\ker\psi}, \mathcal{B}') = \text{Diag}(\alpha_2, \dots, \alpha_n)$$

Comme par construction $\varphi(e_1, e_i) = 0$ pour $i \geqslant 2$, la première ligne de la matrice de Φ dans $\mathcal{B} = \{e_1, \dots, e_n\}$ est $(\alpha_1, 0, \dots, 0)$, et donc

$$\mathcal{M}(\Phi, \mathcal{B}) = \text{Diag}(\alpha_1, \dots, \alpha_n) \quad \blacksquare$$

L'intérêt de la connaissance d'une telle base est évident : en calculant dans cette base, on a alors

$$\Phi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{r} \alpha_i x_i^2$$

Dans la base $\mathcal{B}$, le polynôme homogène représentant Φ ne contient que r termes carrés. Par polarisation, on obtient immédiatement

$$\varphi\left(\sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} y_i e_i\right) = \sum_{i=1}^{r} \alpha_i x_i y_i$$

On dit qu'une telle base réduit la forme quadratique à une forme diagonale. On dit aussi quelque fois que cette base "diagonalise" la forme quadratique, mais cette terminologie est malheureuse et peut entraîner des confusions : si M est,

par exemple, la matrice d'une forme quadratique Φ dans la base canonique de $\mathbb{R}^n$, trouver une base qui réduit Φ, c'est trouver une matrice de passage P transformant M en une matrice diagonale **par congruence**, et en général pas par similitude ; il ne s'agit donc pas de diagonaliser[8] M.

COROLLAIRE 13-3.2 Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est symétrique de rang r si et seulement si M est congruente à une matrice diagonale de rang r.

REMARQUE 13-3.3 Une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ dans laquelle la matrice de $\Phi \in \mathcal{Q}(\mathbb{E}_n)$ est diagonale est une base vérifiant

$$\forall i \neq j \quad \varphi(e_i, e_j) = 0$$

Une telle base est appelée base φ-orthogonale[9] (ou Φ-orthogonale).

13-3.2 Ecriture d'une forme quadratique à l'aide de carrés de formes linéaires

PROPOSITION 13-3.4 Une application $\Phi : \mathbb{E}_n \to \mathbb{R}$ est une forme quadratique de rang r si et seulement s'il existe r formes linéaires indépendantes sur $\mathbb{E}_n$ $(l_i)_{1 \leqslant i \leqslant r}$ et r scalaires $(\alpha_i)_{1 \leqslant i \leqslant r}$ tous non nuls tels que

$$\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$$

Démonstration : Si Φ est une forme quadratique de rang r, on sait qu'il existe une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ dans laquelle l'expression de la forme quadratique est

$$\Phi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{r} \alpha_i x_i^2$$

avec les $(\alpha_i)_{1 \leqslant i \leqslant r}$ tous non nuls. Il suffit alors de prendre pour formes linéaires $(l_i)_{1 \leqslant i \leqslant r}$ les r premières formes coordonnées dans $\mathcal{B}$.

Réciproquement, si $\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$ avec les $(l_i)_{1 \leqslant i \leqslant r}$ indépendantes, on peut compléter la famille $(l_i)_{1 \leqslant i \leqslant r}$ en une base $\mathcal{B}' = (l_i)_{1 \leqslant i \leqslant n}$ de $\mathbb{E}_n^*$. On sait qu'il existe une unique base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de $\mathbb{E}_n$ dont $\mathcal{B}'$ soit la base duale. On a alors

$$\Phi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{r} \alpha_i x_i^2$$

---

[8] La démonstration précédente montre que n'importe quel vecteur e_1 tel que $\Phi(e_1) \neq 0$ peut être choisi comme premier vecteur d'une base réduisant Φ. Nous verrons cependant, dans le chapitre sur les espaces euclidiens, que **dans le cas réel**, une "bonne" diagonalisation de M (c'est à dire avec une matrice de passage pour laquelle ${}^tP = P^{-1}$) est possible.

[9] On dit en effet que deux vecteurs x et y de $\mathbb{E}_n$ sont φ-orthogonaux (ou φ-conjugués) si $\varphi(x, y) = 0$.

ce qui montre que Φ est une forme quadratique[10] dont la matrice dans la base $\mathcal{B}$ est de rang r. ■

13-3.3 Méthode de Gauss

Il s'agit d'une méthode pratique pour écrire un polynôme homogène de degré 2 (donc l'expression d'une forme quadratique dans une base) comme combinaison linéaire de carrés de polynômes homogènes de degré 1 indépendants. On retrouve ici une démonstration "algorithmique" de la proposition **13-3**.4 : on considère un polynôme homogène de degré 2 à n variables (c'est-à-dire une forme quadratique définie sur $\mathbb{R}^n$)

$$\Phi(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{\substack{(i,j) \\ 1 \leqslant i < j \leqslant n}} a_{ij} x_i x_j$$

On commence la réduction en distinguant deux cas :

- **S'il existe des termes carrés :** en renumérotant les variables, on peut supposer que $a_{1,1} \neq 0$. On isole alors dans $\Phi(x_1, \ldots, x_n)$ tous les termes contenant la variable x_1.

$$\Phi(x_1, \ldots, x_n) = a_{1,1} x_1^2 + 2 \sum_{j=2}^{n} a_{1j} x_1 x_j + \Psi(x_2, \ldots, x_n)$$

  où Ψ est une forme quadratique sur $\mathbb{R}^{n-1}$. On écrit alors les premiers termes

[10] Le fait que $\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$ soit une forme quadratique est évident, puisque

$$\varphi(x, y) = \sum_{i=1}^{r} \alpha_i l_i(x) \, l_i(y)$$

est évidemment bilinéaire symétrique associée à Φ (*cf.* la manière dont on polarise un terme carré dans l'écriture d'une forme polaire). Le rang de Φ peut aussi se déterminer en calculant le radical de Φ : pour $x \in \mathbb{E}_n$

$$\varphi(x, \bullet) = 0_{\mathbb{E}_n^*} \Leftrightarrow \sum_{i=1}^{r} \alpha_i l_i(x) \, l_i = 0_{\mathbb{E}_n^*}$$

Comme les $(l_i)_{1 \leqslant i \leqslant r}$ sont supposées indépendantes, ceci équivaut à

$$\forall i \quad \alpha_i l_i(x) = 0$$

soit, puisque les α_i sont non nuls

$$x \in \bigcap_{i=1}^{r} \ker l_i$$

sous-espace de $\mathbb{E}_n$ de dimension $n - r$. Le radical de Φ étant de dimension $n - r$, cette forme quadratique est bien de rang r.

comme le début du développement d'un carré

$$a_{1,1}x_1^2 + 2\sum_{j=2}^{n} a_{1j}x_1x_j = a_{11}\left(x_1 + \sum_{j=2}^{n}\frac{a_{1,j}}{a_{1,1}}x_j\right)^2 + \Psi_1(x_2,\dots,x_n)$$

où $\Psi_1 \in \mathcal{Q}(\mathbb{R}^{n-1})$. En définissant la forme linéaire l_1 sur $\mathbb{R}^n$ par

$$l_1(x_1,\dots,x_n) = x_1 + \sum_{j=2}^{n}\frac{a_{1,j}}{a_{1,1}}x_j$$

et en posant $\Phi_1 = \Psi + \Psi_1$, on obtient

$$\Phi(x_1,\dots,x_n) = a_{11}l_1^2(x_1,\dots,x_n) + \Phi_1(x_2,\dots,x_n)$$

- **Si tous les termes carrés sont nuls** : si Φ n'est pas identiquement nulle, il existe un coefficient $a_{i,j}$ non nul (avec $i \neq j$). En renumérotant les variables, on suppose $a_{1,2} \neq 0$. On isole alors tous les termes contenant x_1 ou x_2 :

$$\Phi(x_1,\dots,x_n) = 2a_{1,2}x_1x_2 + 2\sum_{j=3}^{n} a_{1,j}x_1x_j + 2\sum_{j=3}^{n} a_{2,j}x_2x_j + \Psi(x_3,\dots,x_n)$$

avec cette fois $\Psi \in \mathcal{Q}(\mathbb{R}^{n-2})$. On écrit alors les premiers termes comme le début du développement d'un produit

$$\begin{aligned}&2a_{1,2}x_1x_2 + 2\sum_{j=3}^{n} a_{1,j}x_1x_j + 2\sum_{j=3}^{n} a_{2,j}x_2x_j\\ &\qquad = 2a_{1,2}\left(x_1 + \sum_{j=3}^{n}\frac{a_{2,j}}{a_{1,2}}x_j\right)\left(x_2 + \sum_{j=3}^{n}\frac{a_{1,j}}{a_{1,2}}x_j\right) + \Psi_1(x_3,\dots,x_n)\end{aligned}$$

On définit à présent les formes linéaires

$$\varphi_1(x_1,\dots,x_n) = x_1 + \sum_{j=3}^{n}\frac{a_{2,j}}{a_{1,2}}x_j \quad \text{et} \quad \varphi_2(x_1,\dots,x_n) = x_2 + \sum_{j=3}^{n}\frac{a_{1,j}}{a_{1,2}}x_j$$

et on écrit le produit $\varphi_1\varphi_2$ comme une différence de carrés

$$\varphi_1\varphi_2 = \frac{1}{4}\left[(\varphi_1+\varphi_2)^2 - (\varphi_1-\varphi_2)^2\right]$$

En considérant les formes linéaires $l_1 = \varphi_1 + \varphi_2$ et $l_2 = \varphi_1 - \varphi_2$, on aura

$$\Phi(x_1,\dots,x_n) = \frac{a_{1,2}}{2}l_1^2(x_1,\dots,x_n) - \frac{a_{1,2}}{2}l_2^2(x_1,\dots,x_n) + \Phi_2(x_3,\dots,x_n)$$

avec cette fois $\Phi_2 \in \mathcal{Q}(\mathbb{R}^{n-2})$.

On recommencera ensuite les calculs avec les formes quadratiques Φ_1 ou Φ_2, le nombre de variables ayant diminué d'une ou de deux unités. Le procédé s'arrêtera donc au bout d'un nombre fini d'étapes. Si l'on suppose de plus qu'avec un nombre de variables $\leqslant n-1$, les formes linéaires obtenues par cette méthode sont indépendantes, cette propriété subsistera avec n variables : dans le permier cas, on ajoute aux formes linéaires $(l_i)_{2\leqslant i\leqslant r}$ indépendantes intervenant dans la décomposition de Φ_1 une forme linéaire l_1 qui n'est pas combinaison linéaire des $(l_i)_{2\leqslant i\leqslant r}$, puisque l_1 est seule à faire intervenir la variable x_1. Dans le second cas, si les $(l_i)_{3\leqslant i\leqslant r}$ indépendantes interviennent dans la décomposition de Φ_2, il est clair que $(\varphi_i)_{i=1,2}\cup(l_i)_{3\leqslant i\leqslant r}$ est libre (seule φ_1 fait intervenir x_1 et φ_2 fait intervenir x_2). Il en est de même de la famille équivalente $(l_i)_{1\leqslant i\leqslant r}$.

REMARQUE 13-3.5 La méthode de Gauss laisse le choix de l'ordre dans lequel on fait apparaître les variables : on ne commence pas systématiquement les calculs avec x_1.

EXEMPLE 13-3.6 Trouver une base de $\mathbb{R}^3$ réduisant la forme quadratique dont l'expression, dans la base canonique $\{e_1,e_2,e_3\}$, est

$$\Phi(x,y,z)=x^2+2y^2+2z^2+2xy-2xz+2yz$$

La méthode de Gauss donne

$$\Phi(x,y,z)=(x+y-z)^2+y^2+z^2+4yz$$

soit encore

$$\Phi(x,y,z)=(x+y-z)^2+(y+2z)^2-3z^2$$

La forme Φ est donc de rang[11] 3. Comme les trois[12] formes linéaires

$$\begin{cases} X=x+y-z\\ Y=y+2z\\ Z=z\end{cases}$$

[11]On a donc prouvé par ce calcul que la matrice de Φ dans la base canonique

$$M=\begin{pmatrix}1&1&-1\\1&2&1\\-1&1&2\end{pmatrix}$$

est inversible.

[12]Lorsque le rang de la forme quadratique Φ est inférieur à la dimension de l'espace, on utilise la même méthode, en complétant de manière judicieuse (c'est-à-dire avec des formes linéaires pour lesquelles le calcul sera simple) la famille de formes linéaires $(l_i)_{1\leqslant i\leqslant r}$ intervenant dans la décomposition de Φ en une base $(l_i)_{1\leqslant i\leqslant n}$ de $\mathbb{E}_n^*$. Les égalités

$$\begin{cases} X_1=l_1(x)\\ \quad\vdots\\ X_n=l_n(x)\end{cases}$$

sont alors interprétées comme des formules de changement de bases.

sont indépendantes, on peut les considérer comme les formes coordonnées dans une nouvelle base $\{u_1, u_2, u_3\}$ de $\mathbb{R}^3$. En résolvant ce système, on écrit sa solution sous forme matricielle

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

La matrice intervenant dans cette égalité est matrice de passage de $(e_i)_{1\leqslant i\leqslant 3}$ à $(u_i)_{1\leqslant i\leqslant 3}$. En prenant donc

$$\begin{cases} u_1 = e_1 \\ u_2 = -e_1 + e_2 \\ u_3 = 3e_1 - 2e_2 + e_3 \end{cases}$$

on obtient une base de $\mathbb{R}^3$ dans laquelle la forme Φ s'écrit

$$X^2 + Y^2 - 3Y^2$$

On vérifie d'ailleurs aisément que

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

13-3.4 Signature d'une forme quadratique

Les résultats de cette section sont spécifiques à $\mathbb{R}$. Sur $\mathbb{R}$, les éléments positifs sont exactement les carrés. Lorsqu'on écrit une forme quadratique comme combinaison linéaire de r carrés de formes linéaires indépendantes

$$\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$$

si on suppose que les p premiers coefficients $(\alpha_i)_{1\leqslant i\leqslant p}$ sont strictement positifs et les $r-p$ derniers $(\alpha_i)_{p+1\leqslant i\leqslant r}$ sont strictement négatifs (c'est toujours possible en renumérotant les $(l_i)_{1\leqslant i\leqslant r}$), on aura une écriture encore plus simple de Φ :

$$\Phi = \sum_{i=1}^{p} \varphi_i^2 - \sum_{i=p+1}^{r} \varphi_i^2$$

en posant

$$\begin{aligned} \varphi_i &= \sqrt{\alpha_i} l_i \text{ pour } 1 \leqslant i \leqslant p \\ \varphi_i &= \sqrt{-\alpha_i} l_i \text{ pour } p+1 \leqslant i \leqslant r \end{aligned}$$

les formes linéaires $(\varphi_i)_{1\leqslant i\leqslant r}$ étant toujours évidemment indépendantes. Il existe donc une base de $\mathbb{E}_n$ où la matrice de Φ s'écrit par blocs

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Dans cette base, l'expression de la forme quadratique est

$$\Phi(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2$$

EXERCICE 13-3.7 Retrouver, dans le cas de la dimension finie, le résultat du corollaire **13-1**.18.

EXERCICE 13-3.8 Soient $(l_i)_{1\leqslant i\leqslant m}$ une famille de formes linéaires sur $\mathbb{E}_n$, de rang r. Trouver le rang de la forme quadratique

$$\Phi = \sum_{i=1}^{m} l_i^2$$

On ne peut espérer plus simple que l'écriture $\Phi(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2$. Le théorème qui suit montre que le nombre de coordonnées dont le carré est affecté du signe + ne dépend pas de la méthode suivie pour amener l'écriture de Φ à cette forme réduite :

THÉORÈME 13-3.9 (ET DEFINITION) Soit Φ une forme quadratique de rang r sur un espace réel $\mathbb{E}_n$. Si

$$\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$$

est une écriture de Φ comme combinaison linéaire de carrés de formes linéaires indépendantes, le nombre p de coefficients $(\alpha_i)_{1\leqslant i\leqslant r}$ strictement positifs vaut

$$p = \max\{\dim \mathbb{F} \mid \mathbb{F} \textbf{ s.e.v. de } \mathbb{E}_n \textbf{ tel que } \Phi|_{\mathbb{F}} \textbf{ est définie positive}\}$$

De même, le nombre q de coefficients strictement négatifs est

$$q = \max\{\dim \mathbb{F} \mid \mathbb{F} \textbf{ s.e.v. de } \mathbb{E}_n \textbf{ tel que } \Phi|_{\mathbb{F}} \textbf{ est définie négative}\}$$

Le couple (p, q) a donc une interprétation géométrique, et ne dépend pas de la décomposition de Φ considérée. On l'appelle signature de Φ

$$\operatorname{Sgn}\Phi = (p, q) \textbf{ avec } p + q = \operatorname{rang}\Phi$$

Φ est évidemment positive ssi sa signature est de la forme $(p, 0)$. Elle est définie positive ssi sa signature est $(n, 0)$.

Démonstration : Si l'écriture $\Phi = \sum_{i=1}^{r} \alpha_i l_i^2$ fait intervenir p coefficients > 0 et $q = r - p$ coefficients < 0, nous avons montré dans le préambule de cette section qu'on pouvait trouver une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de $\mathbb{E}_n$ dans laquelle

$$\Phi(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2$$

Si $\mathbb{F}_p = \text{vect}(e_i)_{1 \leqslant i \leqslant p}$, on a $\Phi(x) = \sum_{i=1}^{p} x_i^2$ pour $x \in \mathbb{F}_p$, et il est clair que la restriction de Φ à $\mathbb{F}_p$ est définie positive. On a donc

$$p \leqslant \max\{\dim \mathbb{F} \mid \mathbb{F} \text{ s.e.v. de } \mathbb{E}_n \text{ tel que } \Phi|_{\mathbb{F}} \text{ est définie positive}\}$$

Il y a évidemment égalité si $p = n$. Si $p < n$ et $\mathbb{F}$ est un sous-espace de $\mathbb{E}_n$ de dimension strictement plus grande que p, on a nécessairement

$$\mathbb{F} \cap \text{vect}(e_i)_{p+1 \leqslant i \leqslant n} \neq \{0_{\mathbb{E}_n}\}$$

(les deux sous-espaces ne peuvent être indépendants, puisque la somme de leurs dimensions est strictement supérieure à n). Si x est un vecteur non nul de cette intersection, on a

$$\Phi(x) = -\sum_{i=p+1}^{r} x_i^2 \leqslant 0$$

et la restriction de Φ à $\mathbb{F}$ ne peut être définie positive. Donc

$$p \geqslant \max\{\dim \mathbb{F} \mid \mathbb{F} \text{ s.e.v. de } \mathbb{E}_n \text{ tel que } \Phi|_{\mathbb{F}} \text{ est définie positive}\}$$

ce qui donne bien la valeur annoncée pour p. En changeant Φ en $-\Phi$, on obtient immédiatement la caractérisation de q. ■

COROLLAIRE 13-3.10 Toute matrice $M \in \mathcal{S}_n(\mathbb{R})$ est donc congruente à une matrice canonique (unique) de la forme

$$S_{p,q}^n = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Démonstration : On considère la forme quadratique définie sur $\mathbb{R}^n$ par

$$\Phi(X) = {}^tXMX$$

Si (p, q) est la signature de cette forme quadratique, on sait que M est congruente à $S_{p,q}^n$. Réciproquement, si M est congruente à une matrice $S_{p,q}^n$, la signature de Φ est nécessairement (p, q). ■

EXERCICE 13-3.11 Montrer qu'il y a exactement $\frac{(n+1)(n+2)}{2}$ classes de congruence dans $\mathcal{S}_n(\mathbb{R})$.

EXERCICE 13-3.12 Montrer que dans $\mathcal{S}_n(\mathbb{C})$, il y a exactement $(n+1)$ classes de congruence.

13-3.5 Matrices symétriques positives, définies positives

DÉFINITION 13-3.13 *Si M est une matrice de $\mathcal{S}_n(\mathbb{R})$, on appelle forme quadratique canoniquement associée à M la forme quadratique*[13] *définie sur $\mathbb{R}^n$ par*

$$\Phi_M(X) = {}^tXMX$$

Sa matrice dans la base canonique de $\mathbb{R}^n$ est donc la matrice M. On dit que M est positive (resp. définie positive) ssi $\Phi_M \in Q^+(\mathbb{R}^n)$ (resp. $\Phi_M \in Q^{*+}(\mathbb{R}^n)$). On note $\mathcal{S}_n^+(\mathbb{R})$ et $\mathcal{S}_n^{*+}(\mathbb{R})$ les sous-ensembles de $\mathcal{S}_n(\mathbb{R})$ formés par ces matrices. Il s'agit de deux cônes convexes de $\mathcal{S}_n(\mathbb{R})$

Pour $M \in \mathcal{S}_n(\mathbb{R})$, on a donc

$$\begin{aligned} M &\in \mathcal{S}_n^+(\mathbb{R}) \Leftrightarrow \forall X \in \mathbb{R}^n \quad {}^tXMX \geqslant 0 \\ M &\in \mathcal{S}_n^{*+}(\mathbb{R}) \Leftrightarrow \forall X \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\} \quad {}^tXMX > 0 \end{aligned}$$

REMARQUE 13-3.14 Les coefficients diagonaux d'une matrice $M \in \mathcal{S}_n(\mathbb{R})$ sont les images par Φ_M des vecteurs de la base canonique de $\mathbb{R}^n$. Une matrice positive a donc ses coefficients diagonaux positifs. Ceux d'une matrice de $\mathcal{S}_n^{*+}(\mathbb{R})$ sont strictement positifs. Il n'y a évidemment pas de réciproque, comme le montre l'exemple **13-3**.6.

L'étude de la signature faite à la section précédente donne

PROPOSITION 13-3.15 Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est symétrique positive si et seulement si elle est congruente à une matrice de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

$$\exists P \in \mathcal{GL}_n(\mathbb{R}) \quad M = {}^tP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$$

et M est définie positive si et seulement si M est congruente à I_n

$$\exists P \in \mathcal{GL}_n(\mathbb{R}) \quad M = {}^tPP$$

REMARQUE 13-3.16 Une matrice définie positive est une matrice positive et inversible. Son déterminant est strictement positif.

EXERCICE 13-3.17 Montrer que toute matrice symétrique positive est limite d'une suite de matrices définies positives. Ce résultat est à la base de nombreux raisonnements par densité.

EXERCICE 13-3.18 Montrer que $M \in \mathcal{M}_n(\mathbb{R})$ est symétrique positive si et seulement si elle admet une écriture sous la forme

$$M = {}^tQQ$$

avec $Q \in \mathcal{M}_n(\mathbb{R})$ (non inversible si M n'est pas définie positive).

[13] Φ_M détermine M entièrement. On utilisera donc l'expression "signature de M" pour "signature de Φ_M".

EXERCICE 13-3.19 **Matrices symétriques de rang 1** : montrer qu'une matrice M de $\mathcal{M}_n(\mathbb{R})$ est symétrique positive de rang 1 si et seulement s'il existe un vecteur ligne $L \in \mathbb{R}^n$ non nul avec

$$M = {}^tLL$$

Caractériser, à l'aide de combinaisons linéaires de telles matrices, les matrices symétriques de rang r. Comment obtient-on la signature de M ?

EXERCICE 13-3.20 Montrer que, si $A = (a_{i,j})$ et $B = (b_{i,j})$ sont deux matrices symétriques positives, il en est de même de la matrice $C = (c_{ij})$, avec $c_{ij} = a_{ij}b_{ij}$. (Indication : étudier d'abord le cas où B est de rang 1, en utilisant l'exercice précédent). Que peut-on dire si A et B sont définies positives ?

EXERCICE 13-3.21 Montrer que $\mathcal{S}_n^{*+}(\mathbb{R})$ est un ouvert de $\mathcal{S}_n(\mathbb{R})$. Indication : montrer que son complémentaire est fermé, en munissant $\mathbb{R}^n$ d'une norme quelconque, et en remarquant que

$$M \in \mathcal{S}_n(\mathbb{R}) - \mathcal{S}_n^{*+}(\mathbb{R}) \Leftrightarrow \exists X \in \mathbb{R}^n \quad \|X\| = 1 \text{ et } {}^tXMX \leqslant 0$$

EXERCICE 13-3.22 Montrer qu'une matrice $M \in \mathcal{S}_n(\mathbb{R})$ est définie positive si et seulement si ses mineurs principaux sont tous strictement positifs.

13-3.6 Bases orthonormales pour une forme quadratique définie positive

DÉFINITION 13-3.23 *Si $\Phi \in Q^{*+}(\mathbb{E}_n)$ est un forme définie positive dont φ est la forme polaire, on appelle base Φ-orthonormale*[14] *de $\mathbb{E}_n$ toute famille $(e_i)_{1\leqslant i\leqslant n}$ de n vecteurs de $\mathbb{E}_n$ vérifiant*

$$\forall i \quad \Phi(e_i) = 1 \quad \text{et} \quad \forall i \neq j \quad \varphi(e_i, e_j) = 0$$

Une telle famille est évidemment une base de $\mathbb{E}_n$: comme elle est de cardinal n, il suffit de montrer qu'elle est libre. Or, si $(x_i)_{1\leqslant i\leqslant n}$ sont n réels, on a

$$\sum_{i=1}^n x_ie_i = 0_{\mathbb{E}_n} \Rightarrow \Phi\left(\sum_{i=1}^n x_ie_i\right) = \sum_{i=1}^n\sum_{j=1}^n x_ix_j\varphi(e_i, e_j) = \sum_{i=1}^n x_i^2 = 0 \Rightarrow \forall i \quad x_i = 0$$

Il résulte aussi de ce calcul qu'une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de $\mathbb{E}_n$ est une base Φ-orthonormale si et seulement si l'expression de Φ dans cette base est

$$\Phi(x) = \sum_{i=1}^n x_i^2 \quad \text{(somme des carrés des coordonnées de } x \text{ dans } \mathcal{B})$$

ou si, avec des notations évidentes

$$\varphi(x, y) = \sum_{i=1}^n x_iy_i$$

On déduit immédiatement des résultats sur la réduction des formes quadratiques et la signature le théorème :

[14] Ou simplement base orthonormale s'il n'y a pas d'ambiguïté sur Φ.

THÉORÈME 13-3.24 Si Φ est une forme quadratique définie positive sur $\mathbb{E}_n$, il existe dans cet espace des bases Φ-orthonormales.

Pour le moment, la méthode de Gauss (en suivant la méthode décrite à l'exemple **13-3.6**) est un moyen de prouver qu'une forme quadratique, en dimension finie, est définie positive, et de déterminer des bases orthonormales de l'espace :

EXERCICE 13-3.25 A quelle condition sur le paramètre t la forme quadratique définie sur $\mathbb{R}^3$ par

$$\Phi(x, y, z) = tx^2 + 5y^2 + z^2 - 2xy - 2xz + 4yz$$

est-elle définie positive ? Déterminer une base orthonormale de $\mathbb{R}^3$ dans le cas particulier $t = 5$.

Dans le chapitre qui suit, nous verrons une méthode plus "géométrique" d'obtention de bases orthonormales : le procédé de Gram-Schmidt.

13-4 Exercices

EXERCICE 13-4.1 Soit q la forme quadratique définie sur $\mathbb{R}^4$ par

$$q(X) = x^2 + y^2 + z^2 + t^2 + 2x(y + z + t) - 2y(z + t) - 2zt$$

Ecrire la matrice A de q dans la base canonique. q est-elle dégénérée ? Décomposer q en carrés et calculer A^n.

EXERCICE 13-4.2 Soit $U \in \mathbb{R}^n$ et $\alpha \in \mathbb{R}$. Quelle est la signature de la forme quadratique définie sur $\mathbb{R}^n$ par $\Phi(X) = {}^tXX + \alpha({}^tUX)^2$?

EXERCICE 13-4.3 Sur $\mathbb{R}_n[X]$ pour $k \in 0, \cdots, n$ fixé et $\alpha \in \mathbb{R}$, montrer que Φ définie par

$$\Phi(P) = \int_0^1 \left[P^{(k)}(t)\right]^2 dt + \alpha \sum_{i=0}^{k-1} \left[P^{(i)}(0)\right]^2$$

est une forme quadratique. Quelle est sa signature ?

EXERCICE 13-4.4 Quel est le rang de la matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ avec

$$a_{i,j} = \cos(i + j - 2)\theta$$

EXERCICE 13-4.5 Montrer que la forme quadratique Q de matrice $A = (a_{i,j})$ où $a_{i,j} = i + j - 1$ est le produit de deux formes linéaires.

EXERCICE 13-4.6 Soient $(\varphi_i)_{1 \leqslant i \leqslant p}$ p formes linéaires sur $\mathbb{R}^n$. Quel est le rang de la forme quadratique $\sum \varphi_i^2$?

EXERCICE 13-4.7 Soit E un $\mathbb{R}$-ev de dimension n et φ une forme bilinéaire symétrique sur E de signature $(n-1, 1)$.

1. Soit F un sev de E tel qu'il existe $v \in F$ tel que $\varphi(v,v) < 0$. Quelle est la signature de $\varphi|_{F\times F}$?

2. En déduire une méthode d'obtention des hyperplans H de E tels que $\varphi|_{H\times H}$ soit un produit scalaire.
Etudier en particulier le cas $E = \mathbb{R}^3$ et $\varphi(X,X) = x^2 + 4z^2 + 4xz + 2xy + 2yz$.

EXERCICE 13-4.8 Soit $\alpha \in \mathbb{R}$. Déterminer la signature de la forme quadratique sur $\mathbb{R}^n$

$$\Phi(X) = \sum_{i=1}^{n} x_i^2 + 2\alpha \sum_{1\leqslant i<j\leqslant n} x_i x_j$$

EXERCICE 13-4.9 Déterminer la signature de la forme quadratique sur $\mathbb{R}^{2n}$ de matrice $\begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$.

EXERCICE 13-4.10 Soit Φ une forme quadratique non nulle sur $\mathcal{M}_n(\mathbb{R})$ vérifiant

$$\forall\ X,Y \in \mathcal{M}_n(\mathbb{R}) \quad \Phi(XY) = \Phi(X)\Phi(Y)$$

Montrer que X est inversible ssi $\Phi(X) \neq 0$. Montrer qu'alors $n = 2$ et que, pour tout X de $\mathcal{M}_n(\mathbb{R})$, on a $\Phi(X) = \det(X)$. Ecrire la matrice de Φ dans la base canonique de $\mathcal{M}_2(\mathbb{R})$. Quelle est sa signature ? Soit f un endomorphisme de $\mathcal{M}_2(\mathbb{R})$ et T sa matrice dans la base canonique. Donner une condition nécessaire et suffisante sur T pour que f conserve le déterminant. Donner des exemples de tels endomorphismes. Montrer qu'ils forment un sous-groupe de $\mathcal{GL}(\mathcal{M}_2(\mathbb{R}))$ engendré par trois éléments.

Chapitre 14

Espaces préhilbertiens réels. Espaces euclidiens

14-1 Généralités

14-1.1 Espace préhilbertien

14-1.1.1 Définitions

DÉFINITION 14-1.1 *Un espace préhilbertien réel est un couple* $(\mathbb{E}, \Phi)$*, où* $\mathbb{E}$ *est un* $\mathbb{R}$*-espace vectoriel et* Φ *est une forme quadratique définie positive sur* $\mathbb{E}$.

Sur un $\mathbb{R}$-e.v. $\mathbb{E}$, on peut donc être amené à considérer plusieurs structures d'espace préhilbertien réel : c'est le choix d'une forme quadratique définie positive Φ sur $\mathbb{E}$ qui munit l'espace d'une structure préhilbertienne. Se donner Φ revient aussi à considérer sa forme polaire[1] φ. On dit alors que φ est un produit scalaire sur $\mathbb{E}$:

DÉFINITION 14-1.2 *Si* $\mathbb{E}$ *est un* $\mathbb{R}$*-e.v, un produit scalaire sur* $\mathbb{E}$ *est une forme bilinéaire symétrique sur* $\mathbb{E}$ *dont la forme quadratique associée est définie positive.*

EXEMPLE 14-1.3 Sur $\mathbb{R}^3$, les applications

$$\varphi_1 : \ ((x,y,z),(x',y',z')) \mapsto xx' + yy' + zz'$$

[1]Et on confond les couples $(\mathbb{E}, \Phi)$ et $(\mathbb{E}, \varphi)$.

$$\varphi_2 : ((x,y,z),(x',y',z')) \mapsto (x+y+z)(x'+y'+z') + 2yy' + zz'$$

sont deux produits scalaires.

EXEMPLE 14-1.4 Sur $\mathbb{R}^n$, le produit scalaire canonique est l'application

$$(x,y) = ((x_1,\dots,x_n),(y_1,\dots,y_n)) \mapsto \sum_{i=1}^{n} x_i y_i$$

EXEMPLE 14-1.5 Une généralisation de l'exemple précédent à la dimension infinie est l'espace $l^2(\mathbb{N})$ des suites réelles $u = (u_n)_{n\in\mathbb{N}}$ de carré sommable (*cf.* section **9-3.5**) muni du produit scalaire

$$\langle u, v\rangle = \sum_{n=0}^{+\infty} u_n v_n$$

EXEMPLE 14-1.6 Sur $\mathbf{E} = C^0([a,b],\mathbb{R})$, l'application

$$(f,g) \mapsto \int_a^b f(t)\,g(t)\,dt$$

est un produit scalaire.

EXERCICE 14-1.7 Donner une condition nécessaire et suffisante sur $\omega \in C^0([a,b],\mathbb{R})$ pour que

$$(f,g) \mapsto \int_a^b f(t)\,g(t)\,\omega(t)\,dt$$

définisse un produit scalaire sur l'espace des fonctions réelles définies et continues sur $[a,b]$.

Une **structure préhilbertienne** $(\mathbb{E},\Phi)$ est souvent appelée également **structure euclidienne**, avec un petit abus de langage, puisque le terme "euclidien" devrait être réservé au cas où l'espace $\mathbb{E}$ est de dimension finie.

14-1.1.2 Notations

Lorsqu'une structure préhilbertienne $(\mathbb{E},\varphi)$ est fixée, on note souvent

$$\varphi(x,y) = \langle x,y\rangle \text{ ou } \varphi(x,y) = (x|y)$$

le produit scalaire de deux éléments x et y de $\mathbb{E}$. Ces notations doivent bien sûr être adaptées lorsqu'on considère simultanément plusieurs structures préhilbertiennes sur un même espace $\mathbb{E}$.

Avec la première notation, le carré scalaire d'un vecteur $x \in \mathbf{E}$ sera $\Phi(x) = \langle x,x\rangle$. On le notera le plus souvent $\Phi(x) = \|x\|^2$conformément à ce qui suit :

14-1.1.3 Norme euclidienne

Comme nous l'avons vu à la section **13-1.2.2**, si $(\mathbb{E}, \langle\ ,\ \rangle)$ est un espace préhilbertien réel, l'application

$$\mathbf{E} \to \mathbb{R}^+ \quad x \mapsto \|x\| = \sqrt{\langle x, x\rangle}$$

est une norme, dite norme (euclidienne) associée au produit scalaire sur l'espace $\mathbb{E}$. Plus généralement, on dit qu'une norme N sur l'espace réel $\mathbf{E}$ est une **norme** euclidienne si et seulement si l'application $x \mapsto N^2(x)$ est une forme quadratique (évidemment définie positive) sur $\mathbb{E}$.

L'inégalité de Cauchy-Schwarz s'écrit alors

$$\forall x, y \in \mathbf{E} \quad |\langle x, y\rangle| \leqslant \|x\| \, \|y\|$$

(avec égalité ssi x et y sont liés). Cette inégalité entraîne **la continuité du produit scalaire**, application bilinéaire de $\mathbb{E} \times \mathbb{E}$ dans $\mathbb{R}$.

L'inégalité de Minkowski est simplement l'inégalité triangulaire

$$\forall x, y \in \mathbb{E} \quad \|x + y\| \leqslant \|x\| + \|y\|$$

avec égalité ssi x et y sont colinéaires et de même sens. Les calculs sur les formes quadratiques donnent en particulier

$$\forall x, y \in \mathbb{E} \quad \forall t \in \mathbf{R} \quad \|tx + y\|^2 = t^2 \|x\|^2 + 2t \langle x, y\rangle + \|y\|^2$$

dont on a déduit les formules de "polarisation"

$$\forall x, y \in \mathbb{E} \quad \langle x, y\rangle = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right) = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$$

L'identité qui suit est elle aussi conséquence immédiate de la bilinéarité du produit scalaire :

PROPOSITION 14-1.8 Si $(\mathbb{E}, \langle\ ,\ \rangle)$ est un espace préhilbertien et si $\|\ \|$ représente la norme associée au produit scalaire, on a

$$\forall x, y \in \mathbb{E} \quad \|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

C'est **l'identité du parallélogramme**[2], qui peut être utilisée notamment pour prouver un théorème de projection dans un espace préhilbertien : voir à ce sujet l'exercice **14-2**.15.

---

[2]Dans un espace affine euclidien, si $ABCD$ est un parallélogramme ($\overrightarrow{AB} = \overrightarrow{DC}$) avec

$$\vec{x} = \overrightarrow{AB} \text{ et } \vec{y} = \overrightarrow{AD}$$

on a

$$\vec{x} + \vec{y} = \overrightarrow{AC} \text{ et } \vec{x} - \vec{y} = \overrightarrow{DB}$$

et cette identité montre que, dans un parallélogramme, la somme des carrés des longueurs des diagonales est égale à la somme des carrés des longueurs des quatre côtés.

DÉFINITION 14-1.9 *On appelle espace de Hilbert (réel) tout espace préhilbertien réel* $(\mathbb{E}, \langle\ ,\ \rangle)$ *complet pour la norme associée au produit scalaire.*

En particulier, si l'espace $\mathbb{E}$ est de dimension finie, toute structure préhilbertienne de $\mathbb{E}$ en fait un espace de Hilbert. L'intérêt de la notion d'espace de Hilbert provient du fait que beaucoup de résultats valables dans le cas de la dimension finie sont susceptibles de généralisation à ce type d'espaces. Ils permettent de travailler dans un espace normé (à l'aide d'un produit scalaire) avec un support "visuel" extrêmement utile : celui de la géométrie "euclidienne usuelle" en dimension 2 ou 3.

14-1.2 Orthogonalité

14-1.2.1 Définition

DÉFINITION 14-1.10 *Dans un espace* $(\mathbb{E}, \langle\ ,\ \rangle)$ *préhilbertien réel, deux vecteurs* x *et* y *sont dits orthogonaux (on note alors* $x \perp y$*) si et seulement si*

$$\langle x, y\rangle = 0$$

PROPOSITION 14-1.11 (relation de Pythagore) Deux vecteurs x et y d'un espace préhilbertien sont orthogonaux si et seulement si

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

La relation d'orthogonalité est évidemment symétrique. Plus généralement, si A et B sont deux parties non vides de $\mathbb{E}$ (ou deux familles de vecteurs de $\mathbb{E}$), on dit que A est orthogonale à B si et seulement si tout vecteur de A est orthogonal à tout vecteur de B :

$$A \perp B \Leftrightarrow \forall a \in A \quad \forall b \in B \quad \langle a, b\rangle = 0$$

Par bilinéarité, on a évidemment

PROPOSITION 14-1.12 Si A et B sont deux parties non vides de $\mathbb{E}$

$$A \perp B \Leftrightarrow \operatorname{vect} A \perp \operatorname{vect} B$$

ce qui fait que l'on se ramène souvent à travailler avec des sous-espaces de $\mathbb{E}$ lorsqu'on étudie l'orthogonalité de deux parties de $\mathbb{E}$.

Comme le produit scalaire est défini positif, seul le vecteur nul est orthogonal à lui-même, et cette propriété sera très souvent utilisée pour prouver des égalités dans un espace préhilbertien :

PROPOSITION 14-1.13 Si x et y sont deux vecteurs d'un espace préhilbertien

$$x = y \Leftrightarrow (x-y) \perp (x-y) \Leftrightarrow \{x-y\} \perp \mathbb{E}$$

ce qui peut encore s'écrire

$$x = y \Leftrightarrow \langle x, \bullet\rangle = \langle y, \bullet\rangle \Leftrightarrow \forall z \in \mathbb{E} \quad \langle x, z\rangle = \langle y, z\rangle$$

On peut dire la même chose avec un vocabulaire différent :

PROPOSITION 14-1.14 Si $(\mathbb{E}, \langle\ ,\ \rangle)$ est un espace préhilbertien, l'application $x \mapsto \langle x, \bullet\rangle$ est un morphisme injectif de l'espace $\mathbb{E}$ dans son dual $\mathbb{E}^*$.

14-1.2.2 Famille orthogonale

DÉFINITION 14-1.15 *Une famille $\mathcal{F} = (e_i)_{i\in I}$ de vecteurs d'un espace $(\mathbb{E}, \langle\ ,\ \rangle)$ préhilbertien est orthogonale si et seulement si*

$$\forall i, j \in I \quad i \neq j \Rightarrow e_i \perp e_j$$

PROPOSITION 14-1.16 Une famille orthogonale $\mathcal{F} = (e_i)_{i\in I}$ de vecteurs non nuls d'un espace $(\mathbb{E}, \langle\ ,\ \rangle)$ préhilbertien est libre.

Démonstration : On l'a déjà vu à la section **13-3.6** : si

$$\sum_{i\in I} \alpha_i e_i = 0_{\mathbb{E}}$$

est une relation de liaison entre vecteurs de $\mathcal{F}$, on a

$$\forall j \in I \quad \left\langle \sum_{i\in I} \alpha_i e_i, e_j \right\rangle = \alpha_j \|e_j\|^2 = 0$$

ce qui donne bien $\alpha_j = 0$ puisque e_j est supposé non nul. ■

EXEMPLE 14-1.17 Dans l'espace des fonctions continues $\mathbb{E} = \mathcal{C}^0([0, 2\pi], \mathbf{R})$ si les familles $\mathcal{C} = (c_n)_{n\in\mathbb{N}}$ et $\mathcal{S} = (s_n)_{n\in\mathbb{N}^*}$ sont définies par

$$\forall t \in [0, 2\pi] \quad c_n(t) = \cos nt \quad \text{et} \quad s_n(t) = \sin nt$$

leur réunion est une famille libre. On peut remarquer en effet que $\mathcal{C} \cup \mathcal{S}$ est une famille orthogonale si on munit l'espace $\mathbb{E}$ du produit scalaire

$$\langle f, g\rangle = \int_0^{2\pi} f(t)\, g(t)\, dt$$

C'est un exemple sur lequel nous reviendrons dans le cadre des séries de Fourier.

EXERCICE 14-1.18 Montrer que, plus généralement, si p sous-espaces $(\mathbf{F}_i)_{1\leqslant i\leqslant p}$ d'un espace préhilbertien vérifient

$$i \neq j \Rightarrow \mathbf{F}_i \perp \mathbf{F}_j$$

alors ces p sous-espaces sont indépendants.

14-1.2.3 Orthogonal d'un sous-espace

DÉFINITION 14-1.19 *Si A est une partie non vide d'un espace préhilbertien $(\mathbb{E}, \langle\ ,\ \rangle)$, on appelle orthogonal de A et on note $A^\perp$ l'ensemble des vecteurs orthogonaux à tous les éléments de A :*

$$A^\perp = \{x \in \mathbb{E} \mid \forall a \in A \quad \langle x, a\rangle = 0\}$$

$A^\perp$ est donc la plus grande partie de $\mathbb{E}$ (au sens de l'inclusion) orthogonale à A. C'est clairement un sous-espace vectoriel de $\mathbb{E}$, puisque

$$A^\perp = \bigcap_{a \in A} \ker \langle a, \bullet \rangle$$

est une intersection d'hyperplans (ou l'espace en totalité si A ne contient que le vecteur nul). On a de plus

$$A^\perp = (\mathrm{vect}\, A)^\perp$$

ce qui fait qu'on travaille le plus souvent avec l'orthogonal d'un sous-espace vectoriel[3] de $\mathbb{E}$:

PROPOSITION 14-1.20 Si A et B sont deux parties non vides d'une espace vectoriel préhilbertien

$$A \subset B \Rightarrow B^\perp \subset A^\perp$$

Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}$

$$\mathbb{F} \cap \mathbb{F}^\perp = \{0_\mathbb{E}\} \quad \textbf{et} \quad \mathbb{F} \subset \left(\mathbb{F}^\perp\right)^\perp$$

Si $\mathbb{G}$ est un autre sous-espace vectoriel de $\mathbb{E}$

$$(\mathbb{F} + \mathbb{G})^\perp = \mathbb{F}^\perp \cap \mathbb{G}^\perp \quad \textbf{et} \quad \mathbb{F}^\perp + \mathbb{G}^\perp \subset (\mathbb{F} \cap \mathbb{G})^\perp$$

Démonstration : Si $x \in \mathbb{F} \cap \mathbb{F}^\perp$, il est orthogonal à lui-même, donc nul. Les sous-espaces $\mathbb{F}$ et $\mathbb{F}^\perp$ sont donc indépendants. Les autres propriétés découlent immédiatement de la définition de l'orthogonal d'une partie. ■

REMARQUE 14-1.21 L'inclusion $\mathbb{F} \subset \left(\mathbb{F}^\perp\right)^\perp$ peut être stricte : si on travaille par exemple avec $\mathbb{E} = \mathbb{R}[X]$ muni du produit scalaire

$$\left\langle \sum_{n \in \mathbb{N}} a_n X^n, \sum_{n \in \mathbb{N}} b_n X^n \right\rangle = \sum_{n \in \mathbb{N}} a_n b_n$$

(cette somme est finie), et si

$$\mathbb{F} = \{P \in \mathbb{R}[X] \mid P(1) = 0\}$$

on voit assez facilement que $\mathbb{F}^\perp$ est réduit au polynôme nul (car un polynôme orthogonal à $\mathbb{F}$ est en particulier orthogonal à tout polynôme de la forme $X^m - 1$, où m est un entier naturel quelconque). On a donc $\left(\mathbb{F}^\perp\right)^\perp = \mathbb{R}[X]$, qui contient strictement $\mathbb{F}$. De même, avec $\mathbb{G} = \mathrm{vect}(1)$, on a $\mathbb{F} \cap \mathbb{G} = \{0\}$, donc $(\mathbb{F} \cap \mathbb{G})^\perp = \mathbb{E}$,

[3] L'orthogonal d'un sous-espace vectoriel est donc aussi l'orthogonal d'une famille génératrice ou d'une base de ce sous-espace.

alors que $\mathbb{F}^{\perp} + \mathbb{G}^{\perp}$ est l'hyperplan formé des polynômes qui s'annulent en 0. L'inclusion

$$\mathbb{F}^{\perp} + \mathbb{G}^{\perp} \subset (\mathbb{F} \cap \mathbb{G})^{\perp}$$

est donc stricte. Dans les sections qui suivent, nous verrons que l'égalité

$$\mathbb{F} = \left(\mathbb{F}^{\perp}\right)^{\perp}$$

est toujours vérifiée lorsque l'espace $\mathbb{E}$ est de dimension finie, ou plus généralement lorsque $\mathbb{F}$ est de dimension finie.

EXERCICE 14-1.22 Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}$, montrer que

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp} \Rightarrow \mathbb{F} = \left(\mathbb{F}^{\perp}\right)^{\perp}$$

14-1.3 Cas de la dimension finie

14-1.3.1 Expression du produit scalaire

DÉFINITION 14-1.23 *Un espace vectoriel euclidien est un espace préhilbertien réel de dimension finie.*

Si $(\mathbb{E}_n, \langle\, , \rangle)$ est un tel espace de dimension n, et si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ en est une base, on peut caractériser le produit scalaire par sa matrice dans la base $\mathcal{B}$:

$$A = \mathcal{M}\left(\langle\, , \rangle, \mathcal{B}\right) = \left(\langle e_i, e_j\rangle\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant i \leqslant n}}$$

qui est une matrice symétrique définie positive :

$$\forall X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\} \qquad {}^tXAX = \left\| \sum_{i=1}^{n} x_i e_i \right\|^2 > 0$$

On sait (voir section **13-3.5**) que la matrice A est congruente à la matrice I_n :

$$\exists P \in \mathcal{GL}_n(\mathbb{R}) \quad A = {}^tPP$$

Ceci signifie également qu'il existe une base de $\mathbb{E}_n$, $\mathcal{B}' = (u_i)_{1 \leqslant i \leqslant n}$ telle que

$$\langle u_i, u_j\rangle = 0 \text{ si } i \neq j \text{ et } \langle u_i, u_i\rangle = \|u_i\|^2 = 1$$

Une telle base est une **base orthonormale** de $\mathbb{E}_n$. Dans $\mathcal{B}'$, produit scalaire et norme euclidienne associée ont une expression très simple :

$$\left\| \sum_{i=1}^{n} x_i u_i \right\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

$$\left\langle \sum_{i=1}^{n} x_i u_i, \sum_{i=1}^{n} y_i u_i \right\rangle = \sum_{i=1}^{n} x_i y_i$$

Tout se passe donc comme si l'on calculait dans $\mathbb{R}^n$ muni de sa structure euclidienne canonique. Les coordonnées d'un vecteur **dans une base orthonormale** sont exactement les produits scalaires avec les vecteurs de base, puisque

$$\left\langle \sum_{j=1}^{n} x_j u_j, u_i \right\rangle = x_i$$

On a donc, pour tout vecteur x de $\mathbb{E}_n$ et si la base $(u_i)_{1\leqslant i\leqslant n}$ est orthonormale

$$x = \sum_{i=1}^{n} \langle u_i, x\rangle u_i$$
$$\|x\|^2 = \sum_{i=1}^{n} \langle u_i, x\rangle^2$$

EXERCICE 14-1.24 Montrer que, réciproquement, s'il existe p vecteurs indépendants $(u_i)_{1\leqslant i\leqslant p}$ d'un espace préhilbertien $\mathbb{E}$ tels que

$$\forall x \in \mathbb{E} \quad x = \sum_{i=1}^{p} \langle u_i, x\rangle u_i$$

alors $\mathbb{E}$ est de dimension finie et $(u_i)_{1\leqslant i\leqslant p}$ est une base orthonormale de $\mathbb{E}$.

EXERCICE 14-1.25 Montrer que la matrice $A = (a_{i,j}) \in \mathcal{S}_n(\mathbb{R})$ avec

$$a_{ij} = \frac{1}{i+j}$$

est inversible, puis que son déterminant est strictement positif. On pourra remarquer que

$$\frac{1}{i+j} = \int_0^1 t^{i+j-1}dt$$

14-1.3.2 Isomorphisme entre l'espace euclidien et son dual

Si on fixe une structure euclidienne sur un espace réel $\mathbb{E}_n$ de dimension finie, le produit scalaire est une forme bilinéaire symétrique non dégénérée sur $\mathbb{E}_n$ (car $\langle x, \bullet\rangle = 0_{\mathbb{E}_n^*} \Rightarrow x = 0_{\mathbb{E}_n}$). Il permet donc de définir un isomorphisme "canonique" entre $\mathbb{E}_n$ et son dual. On dit "canonique" car ce morphisme est intrinsèque[4], indépendant en particulier du choix d'une base de $\mathbb{E}_n$:

PROPOSITION 14-1.26 Si $(\mathbb{E}_n, \langle\, ,\, \rangle)$ est un espace euclidien, toute forme linéaire $l \in \mathbb{E}_n^*$ s'écrit de manière unique $\langle z, \bullet\rangle$, avec $z \in \mathbb{E}_n$.

Par l'intermédiaire du produit scalaire, on peut ainsi identifier la forme linéaire l et le vecteur z associé, et donc l'espace $\mathbb{E}_n$ et son dual. Si on veut calculer sur les coordonnées dans une base, cette identification est particulièrement simple si on travaille **dans une base orthonormée** : si $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ est une telle base,

[4] Il dépend cependant de la structure euclidienne fixée sur $\mathbb{E}_n$!

en identifiant un vecteur $x = \sum_{i=1}^{n} x_i e_i$ avec la colonne X de ses coordonnées dans $\mathcal{B}$, une forme linéaire $l \in \mathbb{E}_n^*$ est de la forme

$$X \mapsto \sum_{i=1}^{n} \alpha_i x_i = (\alpha_1, \dots, \alpha_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = LX$$

où L est la ligne $(\alpha_1, \dots, \alpha_n)$. Comme la base est orthonormée, le produit $LX = {}^t({}^tL)X$ peut être interprété comme le produit scalaire $\langle z, x \rangle$, où la colonne représentant les coordonnées de z dans $\mathcal{B}$ est simplement ${}^tL$. On a donc

$$z = \sum_{i=1}^{n} \alpha_i e_i$$

Lorsqu'on travaille dans une base orthonormée, identifier l'espace euclidien et son dual revient tout simplement à transformer les vecteurs lignes en colonnes.

14-1.3.3 Normale à un hyperplan

Si $\mathbb{H}$ est un hyperplan d'un espace euclidien $(\mathbb{E}_n, \langle\ ,\ \rangle)$, on sait qu'il existe une forme linéaire non nulle l sur $\mathbb{E}_n$ (unique à un scalaire multiplicatif non nul près) tel que

$$\mathbb{H} = \ker l$$

Comme il existe un vecteur non nul $x \in \mathbb{E}_n$ tel que $l = \langle x, \bullet \rangle$, on a alors

$$\mathbb{H} = x^{\perp}$$

Le vecteur x engendre une droite vectorielle, appelée **normale** à l'hyperplan $\mathbb{H}$. Comme $x \notin \mathbb{H}$, on a la décomposition de l'espace

$$\mathbb{H} \overset{\perp}{\oplus} \operatorname{vect} x = \mathbb{E}_n$$

(le symbole d'orthogonalité au dessus de l'opération de somme directe signifie que l'on a une décomposition orthogonale de l'espace $\mathbb{E}_n$, notion sur laquelle nous revenons dans le paragraphe qui suit : on a $\mathbb{H} = x^{\perp}$ et on vérifie aisément que $\operatorname{vect} x = \mathbb{H}^{\perp}$). Le vecteur x caractérise l'hyperplan $\mathbb{H}$.

14-1.3.4 Dimension de l'orthogonal d'un sous-espace vectoriel

PROPOSITION 14-1.27 Si $\mathbb{F}$ est un sous-espace vectoriel d'un espace euclidien $\mathbb{E}$

$$\dim \mathbb{F} + \dim \mathbb{F}^{\perp} = \dim \mathbb{E}$$

Démonstration : L'égalité est évidente si $\mathbb{F} = \{0_{\mathbb{E}}\}$. Si $\mathbb{F}$ est de dimension p, et si $(e_i)_{1 \leqslant i \leqslant p}$ est une base de $\mathbb{F}$, les p formes $(\langle e_i, \bullet \rangle)_{1 \leqslant i \leqslant p}$ sont indépendantes (image d'un système libre par l'isomorphisme canonique $\mathbb{E} \to \mathbb{E}^*$). Comme

$$\mathbb{F}^{\perp} = \bigcap_{i=1}^{p} \ker \langle e_i, \bullet \rangle$$

le théorème **3-2**.9 montre que cet espace est de dimension $n - p$, si n est la dimension de $\mathbb{E}$. ■

COROLLAIRE 14-1.28 Si $\mathbb{F}$ est un sous-espace vectoriel d'un espace euclidien, on a

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp}$$

$\mathbb{F}^{\perp}$ est pour cette raison appelé le[5] supplémentaire orthogonal de $\mathbb{F}$.

Démonstration : Les sous-espaces $\mathbb{F}$ et $\mathbb{F}^{\perp}$ sont indépendants, et la somme de leurs dimensions vaut $\dim \mathbb{E}$. ■

COROLLAIRE 14-1.29 Si $\mathbb{F}$ est un sous-espace vectoriel d'un espace euclidien, on a

$$\mathbb{F} = \left(\mathbb{F}^{\perp}\right)^{\perp}$$

Démonstration : On peut utiliser le corollaire précédent et l'exercice **14-1**.22. Il est plus simple d'utiliser l'inclusion et l'égalité des dimensions. ■

En travaillant également par inclusion et égalité des dimensions, on a :

COROLLAIRE 14-1.30 Si $\mathbb{F}$ et $\mathbb{G}$ sont deux sous-espaces vectoriels d'un espace euclidien, on a

$$(\mathbb{F} \cap \mathbb{G})^{\perp} = \mathbb{F}^{\perp} + \mathbb{G}^{\perp}$$

[5] L'usage de l'article défini est conséquence de la propriété suivante, valable dans tout espace préhilbertien (exercice) : si $\mathbb{F}$ et $\mathbb{G}$ sont deux sous-espaces de $\mathbb{E}$,

$$\mathbb{E} = \mathbb{F} \overset{\perp}{\oplus} \mathbb{G} \Rightarrow \mathbb{G} = \mathbb{F}^{\perp}$$

Plus généralement, on dira qu'une décomposition de l'espace $\mathbb{E}$ en somme directe de sous-espaces indépendants

$$\mathbb{E} = \bigoplus_{i=1}^{p} \mathbb{F}_i$$

est une décomposition orthogonale, et on notera alors

$$\mathbb{E} = \overset{\perp}{\bigoplus_{i=1}^{p}} \mathbb{F}_i$$

si et seulement si $\mathbb{F}_i \perp \mathbb{F}_j$ pour $i \neq j$. On vérifie alors aisément que l'on a

$$\forall i \quad \mathbb{F}_i = \left(\bigoplus_{j \neq i} \mathbb{F}_j\right)^{\perp}$$

14-1.3.5 Théorème de la base orthonormale incomplète

Nous savons qu'une famille orthonormale d'un espace préhilbertien est toujours une famille libre. Dans le cas de la dimension finie, une telle famille peut toujours être complétée en une base orthonormale de l'espace :

PROPOSITION 14-1.31 Si $\mathcal{F} = (e_i)_{1 \leqslant i \leqslant p}$ est une famille orthonormale d'un espace euclidien $\mathbb{E}_n$ de dimension n, on peut toujours compléter $\mathcal{F}$ en une base orthonormale de $\mathbb{E}_n$.

Démonstration : Il n'y a rien à faire si $p = n$. Sinon, $\mathbb{F}_p = \operatorname{vect} \mathcal{F}$ est un sous-espace de dimension p de $\mathbb{E}_n$, et on a

$$\mathbb{E}_n = \mathbb{F}_p \oplus \mathbb{F}_p^{\perp}$$

Le sous-espace $\mathbb{F}_p^{\perp}$ (de dimension $n - p$) peut être considéré comme un espace euclidien, lorsqu'on le munit de la restriction du produit scalaire. Les résultats de la section **13-3.5** montrent qu'il possède des bases orthonormales. Si $\mathcal{G} = (u_j)_{1 \leqslant j \leqslant n-p}$ en est une, il est clair que $\mathcal{F} \cup \mathcal{G}$ est une base orthonormale de l'espace $\mathbb{E}_n$. ■

14-2 Projection orthogonale sur un s.e.v. de dimension finie

Lorsque $\mathbb{E}$ est un espace euclidien, nous avons vu que, pour tout sous-espace vectoriel $\mathbb{F}$ de $\mathbb{E}$, on a

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp}$$

Nous savons que cette propriété est fausse[6] en général dans un espace de dimension infinie. Nous allons voir qu'elle est vraie si on suppose que $\mathbb{F}$ est de dimension finie. Pour cela, nous allons commencer par étudier un problème d'approximation :

[6]Une condition **nécessaire** pour que cette égalité soit vraie est que $\mathbb{F}$ soit un sous espace fermé de $\mathbb{E}$ (pour la norme euclidienne). En effet, l'orthogonal d'un sous-espace (non réduit à

14-2.1 Meilleure approximation par un élément d'un sous-espace

Soit $\mathbb{F}$ un sous-espace vectoriel[7] d'un espace préhilbertien réel $\mathbb{E}$ muni de sa norme euclidienne et $a \in \mathbb{E} - \mathbb{F}$. On a défini

$$d(a, \mathbb{F}) = \inf \{\|a - x\|, \ x \in \mathbb{F}\}$$

et on cherche s'il existe un point a_0 dans $\mathbb{F}$ réalisant le minimum de ces distances, c'est-à-dire tel que

$$\|a - a_0\| = d(a, \mathbb{F})$$

En général, la réponse à cette question est négative : si par exemple $\mathbb{F}$ n'est pas fermé[8] dans $(\mathbb{E}, \| \, \|)$, et si l'on prend $a \in \overline{\mathbb{F}} - \mathbb{F}$, on a $d(a, \mathbb{F}) = 0$ alors que tout a_0 dans $\mathbb{F}$ vérifie $\|a - a_0\| > 0$. Un exemple de cette situation est l'espace $\mathbb{E} = \mathcal{C}^0([0,1], \mathbb{R})$ muni du produit scalaire

$$(f, g) = \int_0^1 f(t) g(t) \, dt$$

avec par exemple a égale à la fonction constante égale à 1 et $\mathbb{F}$ l'hyperplan

$$\mathbb{F} = \{f \in \mathbb{E} \mid f(0) = 0\}$$

PROPOSITION 14-2.1 Si $\mathbb{F}$ est un sous-espace vectoriel d'un espace $\mathbb{E}$ préhilbertien et $a \in \mathbb{E} - \mathbb{F}$, il existe dans $\mathbb{F}$ une meilleure approximation de a si et seulement si

$$a \in \mathbb{F} \oplus \mathbb{F}^{\perp}$$

Cette meilleure approximation est alors unique, c'est la composante de a selon $\mathbb{F}$ dans la somme directe $\mathbb{F} \oplus \mathbb{F}^{\perp}$. C'est donc l'unique élément $a_0 \in \mathbb{F}$ tel que $a - a_0 \in \mathbb{F}^{\perp}$.

Démonstration : Supposons d'abord que $a \in \mathbb{F} \oplus \mathbb{F}^{\perp}$, et se décompose dans cette somme directe en $a = a_0 + a_1$. On a alors, pour tout $x \in \mathbb{F}$, compte tenu de la relation de Pythagore

$$\|a - x\|^2 = \|a_1 + (a_0 - x)\|^2 = \|a_1\|^2 + \|a_0 - x\|^2 \geqslant \|a_1\|^2$$

$\{0_{\mathbb{E}}\}$) est intersection d'une famille d'hyperplans fermés. Si $\mathbb{F}$ et $\mathbb{F}^{\perp}$ sont supplémentaires,

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp} \Rightarrow \mathbb{F} = (\mathbb{F}^{\perp})^{\perp}$$

et donc $\mathbb{F}$ est fermé (comme intersection de fermés).

[7]Comme la distance euclidienne est associée à une norme, elle est invariante par translation. L'étude effectuée ici s'appliquera aussi à un sous-espace affine de $\mathbb{E}$: on a tout simplement, pour $\mathbb{F}$ sous-espace vectoriel de $\mathbb{E}$ et $a_0 \in \mathbb{F}$

$$d(a, a_0 + \mathbb{F}) = d(a - a_0, \mathbb{F})$$

et une meilleure approximation de a par un élément de $a_0 + \mathbb{F}$ sera (si elle existe) de la forme $a_0 + a'$, où a' est une meilleure approximation de $a - a_0$ par un élément de $\mathbb{F}$.

[8]Il ne peut exister de sous-espace non fermé que lorsque l'espace $\mathbb{E}$ n'est pas de dimension finie.

avec égalité si et seulement si $x = a_0$. On a donc bien prouvé que

$$d(a, \mathbb{F}) = \|a_1\| = \|a - a_0\|$$

avec unicité du point $x \in \mathbb{F}$ réalisant le minimum de cette distance.

Réciproquement, supposons qu'il existe $a_0 \in \mathbb{F}$ avec

$$d(a, \mathbb{F}) = \|a - a_0\|$$

et montrons que le vecteur $a_1 = a - a_0$ appartient à $\mathbb{F}^{\perp}$, ce qui donnera bien

$$a = a_0 + a_1 \in \mathbb{F} \oplus \mathbb{F}^{\perp}$$

Pour ce faire, choisissons un vecteur $x \in \mathbb{F}$ quelconque et calculons, pour t réel

$$P(t) = \|a - (a_0 + tx)\|^2 = t^2 \|x\|^2 + 2t \langle x, a - a_0 \rangle + \|a - a_0\|^2$$

Cette fonction polynomiale est, par hypothèse, minimale en $t = 0$, et vérifie donc $P'(0) = 0$, ce qui donne

$$\forall x \in \mathbb{F} \quad \langle x, a - a_0 \rangle = 0$$

et on a bien $a - a_0 \in \mathbb{F}^{\perp}$. ■

COROLLAIRE 14-2.2 Lorsque l'espace $\mathbb{E}$ est somme directe de $\mathbb{F}$ et de $\mathbb{F}^{\perp}$, le problème de meilleure approximation d'un élément $a \in \mathbb{E}$ par un élément a_0 de $\mathbb{F}$ possède une solution unique quel que soit a. De plus la correspondance $a \mapsto a_0$ est linéaire : c'est le projecteur sur $\mathbb{F}$ parallèlement à $\mathbb{F}^{\perp}$, qu'on appelle alors projecteur orthogonal sur $\mathbb{F}$.

REMARQUE 14-2.3 Si $a \in \mathbb{F} \oplus \mathbb{F}^{\perp}$, on a aussi $a \in \mathbb{F}^{\perp} \oplus \left(\mathbb{F}^{\perp}\right)^{\perp}$, ce qui montre qu'il existe aussi une meilleure approximation de a par un élément de $\mathbb{F}^{\perp}$: avec les notations précédentes, c'est le vecteur a_1, composante de a selon $\mathbb{F}^{\perp}$. En particulier, le corollaire qui précède et la relation de Pythagore donnent

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp} \Rightarrow \forall a \in \mathbb{E} \quad \|a\|^2 = d^2(a, \mathbb{F}) + d^2\left(a, \mathbb{F}^{\perp}\right)$$

14-2.2 Cas d'un sous-espace de dimension finie

Supposons que $\mathbb{F}$ soit un sous-espace de dimension finie de $\mathbb{E}$. Un argument de **compacité** (valable dans tout espace normé) nous a permis de voir (*cf.* section **7-6.4.1**) que tout élément de $\mathbb{E}$ possédait au moins une meilleure approximation par un élément de $\mathbb{F}$. Comme la norme provient ici d'un produit scalaire, ce qui précède montre l'unicité de cette meilleure approximation. Nous allons ici donner une autre démonstration d'existence, plus "constructive" car elle permet d'expliciter l'approximation :

THÉORÈME 14-2.4 Si $\mathbb{F}$ est un sous-espace de dimension finie d'un espace préhilbertien, on a

$$\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp}$$

Si $(e_i)_{1 \leqslant i \leqslant p}$ est une base orthonormale de $\mathbb{F}$, la meilleure approximation d'un vecteur x par un élément de $\mathbb{F}$ (qui est aussi la projection orthogonale de x sur $\mathbb{F}$) est donnée par

$$x_{\mathbb{F}} = \sum_{i=1}^{p} \langle e_i, x \rangle e_i$$

On a en particulier

$$d^2(x, \mathbb{F}) = \|x - x_{\mathbb{F}}\|^2 \quad \text{et} \quad d^2(x, \mathbb{F}^{\perp}) = \|x_{\mathbb{F}}\|^2 = \sum_{i=1}^{p} \langle e_i, x \rangle^2$$

Démonstration : $\mathbb{F}$ muni de la restriction du produit scalaire est un espace euclidien et possède donc une base orthonormale $(e_i)_{1 \leqslant i \leqslant p}$. Décomposer le vecteur x dans la somme directe $\mathbb{F} \oplus \mathbb{F}^{\perp}$, c'est chercher des scalaires $(\lambda_i)_{1 \leqslant i \leqslant p}$ tels que le vecteur

$$x_{\mathbb{F}} = \sum_{i=1}^{p} \lambda_i e_i$$

vérifie $x - x_{\mathbb{F}} \in \mathbb{F}^{\perp}$, soit

$$\forall j \quad \langle x - x_{\mathbb{F}}, e_j \rangle = 0$$

ce qui donne exactement, puisque la base $(e_i)_{1 \leqslant i \leqslant p}$ est supposée orthonormale

$$\forall j \quad \lambda_j = \langle x, e_j \rangle$$

et démontre[9] le résultat escompté. ■

[9] L'intérêt de disposer d'une base orthonormale de $\mathbb{F}$ est d'obtenir une "formule" permettant de projeter un vecteur x sur $\mathbb{F}$. Pour prouver l'existence d'une décomposition de x dans $\mathbb{F} \oplus \mathbb{F}^{\perp}$, on aurait pu considérer la forme linéaire $\langle x, \bullet \rangle|_{\mathbb{F}}$ sur l'espace euclidien $\mathbb{F}$, associée à un vecteur $x_{\mathbb{F}}$ unique de $\mathbb{F}$ (isomorphisme entre un espace euclidien et son dual). On a ainsi

$$\exists x_{\mathbb{F}} \in \mathbb{F} \quad \forall y \in \mathbb{F} \quad \langle x, y \rangle = \langle x_{\mathbb{F}}, y \rangle$$

et donne bien $x - x_{\mathbb{F}} \in \mathbb{F}^{\perp}$. Le calcul explicite de $x_{\mathbb{F}}$ peut se faire en travaillant dans une base $(u_i)_{1 \leqslant i \leqslant p}$ de $\mathbb{F}$ non nécessairement orthonormale. Si on cherche $x_{\mathbb{F}}$ sous la forme

$$x_{\mathbb{F}} = \sum_{i=1}^{p} \lambda_i u_i$$

les conditions

$$\forall j \quad \langle x - x_{\mathbb{F}}, u_j \rangle = 0$$

COROLLAIRE 14-2.5 **Si $\mathbb{F}$ est un sous-espace de dimension finie d'un espace préhilbertien, $\left(\mathbb{F}^{\perp}\right)^{\perp} = \mathbb{F}$.**

EXEMPLE 14-2.6 **Distance à une droite vectorielle, à l'hyperplan orthogonal** : si a est un vecteur non nul d'un espace préhilbertien, la projection orthogonale d'un vecteur x sur la droite vectorielle $\mathbb{D} = \operatorname{vect} a$ est donnée par

$$x_{\mathbb{D}} = \frac{\langle x, a\rangle}{\|a\|^2} a$$

(avec les notations qui précèdent $p = 1$ et $e_1 = \dfrac{a}{\|a\|}$). La distance de x à la droite $\mathbb{D}$ est donc donnée par

$$d(x, \operatorname{vect} a) = \left\| x - \frac{\langle x, a\rangle}{\|a\|^2} a \right\|$$

alors que la distance à l'hyperplan orthogonal est

$$d\left(x, a^{\perp}\right) = \frac{|\langle x, a\rangle|}{\|a\|}$$

On retrouve le résultat bien connu en géométrie élémentaire : dans l'espace affine euclidien de dimension 3 rapporté à un **repère orthonormé** $\left(O, \vec{I}, \vec{J}, \vec{K}\right)$, la distance d'un point $M(x, y, z)$ à un plan $\mathcal{P}$ d'équation $uX + vY + wZ + h = 0$ est donnée par

$$d(M, \mathcal{P}) = \frac{|ux + vy + wz + h|}{\sqrt{u^2 + v^2 + w^2}}$$

En effet, si $M_0(x_0, y_0, z_0)$ est un point de $\mathcal{P}$ et $\vec{a} = u\vec{I} + v\vec{J} + w\vec{K}$, on a

$$\mathcal{P} = M_0 + (\vec{a})^{\perp}$$

et, par invariance de la distance par translation

$$d(M, \mathcal{P}) = d\left(\overrightarrow{M_0 M}, (\vec{a})^{\perp}\right) = \frac{\left|\overrightarrow{M_0 M}.\vec{a}\right|}{\|\vec{a}\|}$$

et $\overrightarrow{M_0 M}.\vec{a} = \overrightarrow{OM}.\vec{a} - \overrightarrow{OM_0}.\vec{a} = ux + vy + wz + h$, puisque

$$\overrightarrow{OM_0}.\vec{a} = ux_0 + vy_0 + wz_0 = -h$$

amèneront cette fois à résoudre un système

$$(\langle u_i, u_j\rangle) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} \langle x, u_1\rangle \\ \vdots \\ \langle x, u_p\rangle \end{pmatrix}$$

de Cramer, puisque la matrice du produit scalaire dans la base $(u_i)_{1 \leqslant i \leqslant p}$ est inversible.

14-2.3 Orthogonalisation de Gram-Schmidt

L'obtention d'une base orthonormale d'un espace euclidien est possible par différentes méthodes : nous avons déjà vu la méthode de Gauss qui, à partir de l'expression du carré scalaire dans une base quelconque permet d'obtenir une base où ce carré scalaire s'exprime comme somme des carrés des coordonnées. Une telle base est orthonormale.

Des considérations géométriques vont nous donner un algorithme de construction d'une famille orthonormale équivalente à une famille libre donnée d'un espace préhilbertien. Dans le cas d'un espace euclidien, si on part d'une base de l'espace, on obtiendra une base orthonormale :

THÉORÈME 14-2.7 Soit $(u_i)_{1\leqslant i\leqslant p}$ une famille libre d'un espace préhilbertien $\mathbb{E}$. Il existe une famille orthonormale unique $(e_i)_{1\leqslant i\leqslant p}$ de l'espace $\mathbb{E}$ telle que

$$\forall i \in \{1,\dots,p\} \quad \operatorname{vect}(e_1,\dots,e_i) = \operatorname{vect}(u_1,\dots,u_i)$$
$$\forall i \in \{1,\dots,p\} \quad \langle e_i, u_i\rangle > 0$$

Cette famille peut être construite de proche en proche par l'algorithme de Gram-Schmidt :

$$e_1 = \frac{u_1}{\|u_1\|} \textbf{ et pour } i \geqslant 2 \quad e_i = \frac{u'_i}{\|u'i\|} \textbf{ avec } u'_i = u_i - \sum_{k=1}^{i-1} \langle e_k, u_i\rangle\, e_k$$

Démonstration : La première condition traduit le fait que la famille $(e_i)_{1\leqslant i\leqslant p}$ est triangulaire par rapport à $(u_i)_{1\leqslant i\leqslant p}$. Nous verrons dans le courant de la démonstration que la seconde condition est là pour assurer l'unicité de la famille $(e_i)_{1\leqslant i\leqslant p}$: sans cette condition, il y aurait 2^p familles orthonormales répondant à la question.

La démonstration se fait par récurrence sur p. Si $p = 1$, on cherche simplement un vecteur e_1 unitaire et colinéaire à u_1. Il y a deux possibilités

$$e_1 = \pm\frac{u_1}{\|u_1\|}$$

mais la condition de signe portant sur (e_1, u_1) impose le signe $+$. Supposons acquis l'énoncé pour une famille libre de $p-1$ vecteurs de $\mathbb{E}$ et donnons nous une famille libre $(u_i)_{1\leqslant i\leqslant p}$. En appliquant l'hypothèse de récurrence à $(u_i)_{1\leqslant i\leqslant p-1}$, on obtient une base $(e_i)_{1\leqslant i\leqslant p-1}$ orthonormale unique de $\operatorname{vect}(u_i)_{1\leqslant i\leqslant p-1}$ triangulaire par rapport à la famille $(u_i)_{1\leqslant i\leqslant p-1}$ et vérifiant $\langle e_i, u_i\rangle > 0$ pour $i \in \{1,\dots,p-1\}$, donnée par

$$e_1 = \frac{u_1}{\|u_1\|} \text{ et pour } i \geqslant 2 \quad e_i = \frac{u'_i}{\|u'i\|} \text{ avec } u'_i = u_i - \sum_{k=1}^{i-1} \langle e_k, u_i\rangle\, e_k$$

Il reste donc à prouver l'existence et l'unicité du vecteur e_p complétant $(e_i)_{1\leqslant i\leqslant p-1}$ en une base orthonormale de $\mathbb{F}_p = \operatorname{vect}(u_i)_{1\leqslant i\leqslant p}$. Procédons par analyse-synthèse :

Analyse : Dans l'espace euclidien $\mathbb{F}_p$ (pour la restriction du produit scalaire), nous disposons d'une base orthonormale $(e_i)_{1\leqslant i\leqslant p-1}$ de $\mathbb{F}_{p-1} = \text{vect}(u_i)_{1\leqslant i\leqslant p-1}$. On cherche le vecteur $e_p \in \mathbb{F}_p$ orthogonal à $(e_i)_{1\leqslant i\leqslant p-1}$. Il est donc dans la droite vectorielle $\mathbb{F}_p \cap \mathbb{F}_{p-1}^{\perp}$ (droite vectorielle, puisqu'il s'agit de l'orthogonal dans l'espace euclidien $\mathbb{F}_p$ de l'hyperplan $\mathbf{F}_{p-1}$). Pour avoir un vecteur directeur de cette droite, il suffit de prendre un vecteur dans $\mathbb{F}_p - \mathbb{F}_{p-1}$ et de lui retrancher sa projection orthogonale sur $\mathbb{F}_{p-1}$. Comme on dispose du vecteur u_p, on choisira

$$u'_p = u_p - \sum_{k=1}^{p-1} \langle e_k, u_i\rangle\, e_k$$

On a donc nécessairement (figure **14.1**) :

$$e_p = \pm \frac{u'_p}{\|u'_p\|}$$

vecteur unitaire orthogonal à $(e_k)_{1\leqslant k\leqslant p}$. On a alors

$$\begin{aligned}\pm\left\|u'_p\right\| = \langle e_p, u'_p\rangle &= \left\langle e_p, u_p - \sum_{k=1}^{p-1} \langle e_k, u_i\rangle\, e_k\right\rangle \\ &= \langle e_p, u_p\rangle - \left\langle e_p, \sum_{k=1}^{p-1} \langle e_k, u_i\rangle\, e_k\right\rangle = \langle e_p, u_p\rangle\end{aligned}$$

ce qui montre que l'on doit choisir le signe + pour avoir $\langle e_p, u_p\rangle > 0$. (on voit que, sans cette condition de signe, on aurait, à chaque étape de la récurrence, 2 possibilités pour construire le vecteur e_i, ce qui donnerait à l'arrivée 2^p familles orthonormales possibles).

Synthèse : Les calculs précédents montrent que le vecteur e_p ainsi construit est unitaire, orthogonal à $(e_i)_{1\leqslant i\leqslant p-1}$ et vérifie

$$\begin{gathered}\text{vect}(e_1, \ldots, e_p) = \text{vect}(u_1, \ldots, u_p) \\ \langle e_p, u_p\rangle > 0.\end{gathered}$$

et répond donc à la question.

On a donc bien prouvé, à l'ordre p l'existence et l'unicité de la famille $(e_i)_{1\leqslant i\leqslant p}$, qu'on appellera orthonormalisée de Schmidt de la famille $(u_i)_{1\leqslant i\leqslant p}$. ■

REMARQUE 14-2.8 Le fait d'imposer aux vecteurs de la famille $(e_i)_{1\leqslant i\leqslant p}$ d'être unitaires oblige à chaque étape à diviser chaque vecteur u'_i par sa norme, qui s'exprime comme racine carrée d'un carré scalaire, et oblige souvent à traîner des radicaux dans les calculs. On peut préférer l'orthogonalisation de Schmidt, en conservant la famille $(u'_i)_{1\leqslant i\leqslant p}$ qui est cette fois orthogonale, triangulaire par rapport à $(u_i)_{1\leqslant i\leqslant p}$ (et vérifie $\langle u'_i, u_i\rangle > 0$ pour tout i). L'algorithme d'obtention s'écrit alors (ne pas oublier de diviser par les carrés des normes) :

$$u'_1 = u_1 \quad \text{et pour } i \geqslant 2 \quad u'_i = u_i - \sum_{k=1}^{i-1} \langle u'_k, u_i\rangle \frac{u'_k}{\|u'_k\|^2}$$

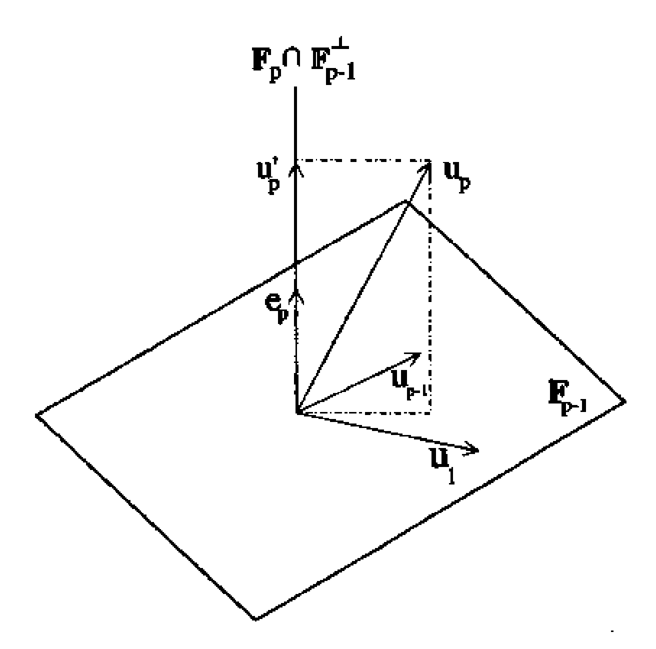


Figure **14.1** – *Procédé d'orthonormalisation de Schmidt*

REMARQUE 14-2.9 Si l'espace $\mathbb{E}$ est de dimension infinie, et si $\mathcal{U} = (u_n)_{n\in\mathbb{N}}$ est une suite de vecteurs libres de $\mathbb{E}$, on peut poursuivre indéfiniment l'algorithme d'orthonormalisation pour obtenir une suite orthonormale $\mathcal{E} = (e_n)_{n\in\mathbb{N}}$ triangulaire par rapport à $\mathcal{U}$.

EXERCICE 14-2.10 Déterminer $\inf_{(a,b)\in\mathbb{R}^2} \int_0^{2\pi} (\sin t + at + b)^2\, dt$

EXERCICE 14-2.11 Quelle est la matrice, dans la base canonique de $\mathbb{R}^4$ muni de sa structure euclidienne canonique, du projecteur orthogonal sur le sous-espace $\mathbf{F}$ d'équations

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \end{cases}$$

EXERCICE 14-2.12 Si $\langle\, ,\rangle$ est un produit scalaire sur l'espace $\mathbb{R}[X]$, montrer qu'il existe une unique suite de polynômes $(P_n)_{n\in\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N} \quad P_n \text{ est un polynôme normalisé de degré } n$$
$$n \neq m \Rightarrow P_n \perp P_m$$

Le procédé d'orthonormalisation de Schmidt a une interprétation matricielle intéressante :

PROPOSITION 14-2.13 Soit $M \in \mathcal{S}_n^{*+}(\mathbb{R})$ une matrice symétrique définie positive. Il existe une unique matrice T triangulaire supérieure réelle, dont les coefficients diagonaux sont strictement positifs, telle que

$$M = {}^tTT$$

Démonstration : L'unicité s'obtient par un raisonnement purement matriciel : si on a deux décompositions

$$M = {}^tTT = {}^tSS$$

l'inversibilité de M entraîne celle de S et T et

$${}^tS^{-1}{}^tT = {}^t\left(TS^{-1}\right) = ST^{-1}$$

égalité entre une matrice triangulaire inférieure et une matrice triangulaire supérieure. Ces deux matrices sont donc diagonales, à coefficients diagonaux strictement positifs et inverses l'une de l'autre. Elles sont donc égales à I_n et donc $ST^{-1} = I_n \Rightarrow S = T$.

Le raisonnement géométrique qui suit donne facilement l'existence de la matrice T (et donnerait aussi l'unicité). On interprète M comme la matrice d'un produit scalaire dans la base canonique $\mathcal{B} = (u_i)_{1\leqslant i\leqslant n}$ de $\mathbb{R}^n$. Le procédé d'orthonormalisation de Schmidt fournit alors une base $\mathcal{B}' = (e_i)_{1\leqslant i\leqslant n}$ orthonormale pour ce produit scalaire, telle que la matrice de passage U de $\mathcal{B}$ à $\mathcal{B}'$ soit triangulaire supérieure à coefficients diagonaux strictement positifs (avec les notations de la démonstration du théorème, ces coefficients diagonaux sont les $\left(\dfrac{1}{\|u_i'\|}\right)_{1\leqslant i\leqslant n}$).
La formule de changement de base pour la matrice d'une forme bilinéaire symétrique donne alors

$${}^tUMU = I_n$$

soit $M = {}^tTT$ avec $T = U^{-1}$, matrice triangulaire supérieure, dont les coefficients diagonaux sont inverses de ceux de U et sont donc aussi strictement positifs. ■

L'intérêt de ce résultat matriciel est évident : si on doit résoudre le système linéaire[10] d'inconnue X

$$MX = B$$

[10] Le fait que la matrice du système soit symétrique définie positive n'est, en théorie, pas restrictif : un système de Cramer de n équations à n inconnues peut s'écrire, en travaillant avec les vecteurs colonnes de la matrice du système

$$x_1C_1 + \cdots + x_nC_n = B$$

et s'interpréter comme la recherche des coordonnées d'un vecteur $B \in \mathbb{R}^n$ dans la base $(C_i)_{1\leqslant i\leqslant n}$. Si on munit $\mathbb{R}^n$ d'une structure euclidienne, par exemple la structure canonique, ce système entraîne

$$\forall i \quad \langle C_i, x_1C_1 + \cdots + x_nC_n\rangle = \langle C_i, B\rangle$$

ce qui donne un système d'équations (dit "normal")

$$(\langle C_i, C_j\rangle)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \langle C_1, B\rangle \\ \vdots \\ \langle C_n, B\rangle \end{pmatrix}$$

conséquence du système de départ, et équivalent à celui-ci : c'est en effet lui-aussi un système de Cramer, puisque sa matrice est définie positive, c'est la matrice du produit scalaire dans la base

et qu'on dispose d'un algorithme de factorisation de M sous la forme ${}^tTT$, le système sera résolu en posant d'abord $Y = TX$ et en résolvant le système triangulaire ${}^tTY = B$, puis en résolvant le système triangulaire $TX = Y$. Remarquons que le procédé de Schmidt appliqué à la base canonique de $\mathbb{R}^n$ (attention ! pour le produit scalaire associé à M) donne la matrice $U = T^{-1}$ (ce qui est encore mieux puisque la solution du système est alors $X = U{}^tUB$). Mais cet algorithme n'est pas utilisé en pratique car il est numériquement instable, et est donc très sensible à la propagation des erreurs d'arrondi.

14-2.4 Inégalité de Bessel

Lorsqu'on travaille sur un espace euclidien $\mathbf{E}_n$ de dimension n rapporté à une base orthonormale $(e_i)_{1\leqslant i\leqslant n}$, on a

$$\forall x \in \mathbf{E}_n \quad \|x\|^2 = \sum_{i=1}^{n} \langle e_i, x\rangle^2$$

Nous allons voir que cette égalité peut être généralisée à certains vecteurs d'un espace préhilbertien de dimension infinie, si on considère une suite orthonormale $(e_i)_{i\in\mathbb{N}}$ (obtenue par exemple à partir d'une suite libre par le procédé de Schmidt). Nous reviendrons longuement sur ce sujet dans le chapitre sur les séries de Fourier.

PROPOSITION 14-2.14 (Inégalité de Bessel) Si $(e_n)_{n\in\mathbb{N}}$ est une suite[11] orthonormale d'un espace préhilbertien, pour tout $x \in \mathbb{E}$, la suite $(\langle e_n, x\rangle)_{n\in\mathbb{N}}$ est de carré sommable et

$$\sum_{n=0}^{+\infty} \langle e_n, x\rangle^2 \leqslant \|x\|^2$$

Il y a égalité dans cette inégalité si et seulement si la suite $(x_n)_{n\in\mathbb{N}}$ des projections orthogonales de x sur les sous-espaces $\mathbb{E}_n = \mathrm{vect}\,(e_i)_{0\leqslant i\leqslant n}$

$$x_n = \sum_{i=0}^{n} \langle e_i, x\rangle\, e_i$$

converge dans l'espace normé $(\mathbb{E}, \|\ \|)$ vers x, ce qui revient exactement à dire que le vecteur x appartient à l'adhérence du sous-espace $\mathrm{vect}\,(e_n)_{n\in\mathbb{N}}$.

$(C_i)_{1\leqslant i\leqslant n}$. Il n'a donc qu'une seule solution, qui est celle du système de départ. (on pourrait aussi dire, par un raisonnement plus géométrique, que le vecteur B et donc ses coordonnées dans la base $(C_i)_{1\leqslant i\leqslant n}$ sont entièrement déterminés par les valeurs que prennent sur lui les n formes linéaires indépendantes $\langle C_i, \bullet\rangle$, donc lorsqu'on connaît ses produits scalaires avec les vecteurs de cette base).

[11]En modifiant légèrement la démonstration (exercice) on obtiendrait l'énoncé :

Si I est un ensemble dénombrable et $(e_i)_{i\in I}$ est une famille orthonormale d'un espace préhilbertien $\mathbb{E}$, pour tout x de $\mathbb{E}$, la famille de réels positifs $\left(\langle x, e_i\rangle^2\right)_{i\in I}$ est sommable et

$$\sum_{i\in I} \langle x, e_i\rangle^2 \leqslant \|x\|^2$$

Quand y a-t-il égalité ?

Démonstration : Si x_n représente la projection orthogonale de x sur $\mathbb{E}_n$, on a

$$x_n = \sum_{i=0}^{n} \langle e_i, x\rangle\, e_i \quad \text{et donc} \quad \|x_n\|^2 = \sum_{i=0}^{n} \langle e_i, x\rangle^2$$

(expression du carré scalaire dans une base orthonormale). La relation de Pythagore appliquée à la décomposition $x = x_n + (x - x_n)$ donne $\|x\|^2 = \|x_n\|^2 + \|x - x_n\|^2$ et donc

$$\|x_n\|^2 = \sum_{i=0}^{n} \langle e_i, x\rangle^2 \leqslant \|x\|^2$$

Les sommes partielles de la séries de terme général positif $\langle e_n, x\rangle^2$ sont donc majorées. Cette série est donc convergente, et

$$\sum_{n=0}^{+\infty} \langle e_n, x\rangle^2 \leqslant \|x\|^2$$

Comme de plus

$$\|x - x_n\|^2 = \|x\|^2 - \|x_n\|^2 = \|x\|^2 - \sum_{i=0}^{n} \langle e_i, x\rangle^2$$

il y a égalité dans l'inégalité précédente si et seulement si

$$\lim_{n\to+\infty} \|x - x_n\| = 0$$

ce qui traduit le fait que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x. Comme

$$\forall n \quad x_n \in \operatorname{vect}(e_i)_{i\in\mathbb{N}}$$

cette convergence entraîne que $x \in \overline{\operatorname{vect}(e_n)_{n\in\mathbb{N}}}$ comme limite d'une suite convergente de points de cet espace. Réciproquement, si un vecteur $x \in \overline{\operatorname{vect}(e_n)_{n\in\mathbb{N}}}$ et si $\varepsilon > 0$ est choisi arbitrairement, on peut trouver un vecteur $y \in \operatorname{vect}(e_n)_{n\in\mathbb{N}}$ vérifiant $\|x - y\| \leqslant \varepsilon$. Comme une combinaison linéaire d'une famille de vecteurs ne fait intervenir qu'une sous-famille finie, il existe un entier m avec $y \in \mathbb{E}_m$. Par propriété de meilleure approximation, on a

$$\|x - x_m\| \leqslant \|x - y\| \leqslant \varepsilon$$

Comme la suite $(\|x - x_n\|)_{n\in\mathbb{N}}$ est décroissante, on a

$$\forall n \geqslant m \quad \|x - x_n\| \leqslant \varepsilon$$

et il y a donc bien convergence de la suite $(x_n)_{n\in\mathbb{N}}$ vers x. ∎

14-2.5 Complément : projection sur un convexe complet

EXERCICE 14-2.15 Si $\mathbf{E}$ est un espace préhilbertien et C est un convexe complet de $\mathbb{E}$ (si $\mathbf{E}$ est un espace de Hilbert, cela revient à supposer que C est convexe fermé). On se propose de montrer que

$$\forall a \in \mathbb{E} \quad \exists!\, b_a \in C \quad \|a - b_a\| = \inf_{x \in C} \|a - x\|$$

b_a est alors appelé projection de a sur C.

Pour cela, montrer que, si a, y et $z \in \mathbf{E}$ et $m = \frac{1}{2}(y + z)$ est le milieu de $[y, z]$

$$\|a - y\|^2 + \|a - z\|^2 = 2\|a - m\|^2 + \frac{1}{2}\|y - z\|^2$$

et en déduire l'unicité d'une projection éventuelle. Montrer ensuite que, si $(x_n)_{n \in \mathbb{N}}$ est une suite de points de C telle que

$$\lim_{n \to +\infty} \|a - x_n\| = d(a, C)$$

cette suite est de Cauchy. Conclure.

Montrer que la projection b_a de a sur C est caractérisée par la propriété

$$b_a \in C \quad \text{et} \quad \forall x \in C \quad \langle b_a - x, b_a - a \rangle \leqslant 0$$

EXERCICE 14-2.16 Soit $(\mathbb{E}, \langle\ ,\ \rangle)$ un espace de Hilbert et φ une forme linéaire non nulle continue sur $\mathbb{E}$. Montrer que $\mathbf{H} = \ker \varphi$ est un hyperplan fermé de $\mathbf{E}$. Si a est un vecteur de $\mathbb{E} - \mathbb{H}$, montrer que a possède une meilleure approximation par un élément de $\mathbb{H}$. En déduire que

$$\mathbf{E} = \mathbb{H} \oplus \mathbb{H}^{\perp}$$

et montrer que $\mathbb{H}^{\perp}$ est une droite vectorielle. Si x en est un vecteur directeur, montrer qu'il existe $\lambda \in \mathbb{R}^*$ avec

$$\varphi = \langle \lambda x, \bullet \rangle$$

Si on note $\mathbf{E}'$ le dual topologique de $\mathbf{E}$, montrer que l'application

$$\mathbf{E} \to \mathbf{E}' \quad z \mapsto \langle z, \bullet \rangle$$

est un isomorphisme d'espace vectoriel qui est une isométrie si on munit $\mathbf{E}'$ de la norme des applications linéaires continues. On voit donc qu'une propriété fondamentale des espaces euclidiens se conserve dans les espaces de Hilbert, si on se limite à travailler avec des formes linéaires continues.

14-3 Endomorphismes d'un espace euclidien

Nous nous intéressons dans cette section aux endomorphismes d'un espace préhilbertien de dimension finie $(\mathbb{E}_n, \langle\ ,\ \rangle)$. Si $u \in \mathcal{L}(\mathbb{E}_n)$, on peut représenter u par sa matrice dans une base arbitraire de l'espace. Si cette base $\mathcal{B} = \{e_1, \ldots, e_n\}$ est **orthonormale**, les coefficients de cette matrice étant obtenus comme coordonnées dans $\mathcal{B}$ des images des vecteurs de $\mathcal{B}$, coordonnées qui s'expriment aussi comme des produits scalaires

$$M(u, \mathcal{B}) = (\langle e_i, u(e_j)\rangle)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(\mathbb{R})$$

On a en particulier

$$\operatorname{trace} u = \sum_{i=1}^{n} \langle e_i, u(e_i)\rangle$$

Nous verrons ultérieurement une interprétation géométrique du déterminant.

14-3.1 Adjoint d'un endomorphisme

14-3.1.1 Définition

THÉORÈME 14-3.1 (ET DEFINITION) Si u est un endomorphisme d'un espace euclidien $(\mathbb{E}_n, \langle\ ,\ \rangle)$, il existe un unique endomorphisme de $\mathbb{E}_n$ appelé adjoint de u et noté u^* tel que

$$\forall x, y \in \mathbb{E}_n \quad \langle u(x), y\rangle = \langle x, u^*(y)\rangle$$

Démonstration : On voit que l'endomorphisme u^* est défini par dualité, c'est-à-dire que l'on connaît l'image $u^*(y)$ d'un vecteur quelconque de $\mathbb{E}_n$ par ses produits scalaires avec tous les vecteurs de l'espace. Nous prouvons donc l'existence et l'unicité de u^* en utilisant l'isomorphisme canonique entre $\mathbb{E}_n$ et $\mathbb{E}_n^*$:

Unicité : celle-ci résulte de la remarque précédente : si $y \in \mathbb{E}_n$ est quelconque, et si u^* existe, on a

$$\langle u^*(y), \bullet\rangle \underset{\text{déf}}{=} \langle y, u(\bullet)\rangle$$

ce qui montre l'unicité de $u^*(y)$, donc celle de u^*.

Existence : Si $y \in \mathbb{E}_n$, l'application $x \mapsto \langle y, u(x)\rangle$ est clairement une forme linéaire sur $\mathbb{E}_n$. Comme le produit scalaire est non-dégénéré, il existe un unique vecteur $z_y \in \mathbb{E}_n$ tel que

$$\forall x \in \mathbb{E}_n \quad \langle z_y, x\rangle = \langle y, u(x)\rangle$$

Notons u^* l'application $y \mapsto u^*(y) = z_y$ de $\mathbb{E}_n$ dans lui-même. Si y_1 et y_2 sont deux vecteurs de $\mathbb{E}_n$ et $\alpha \in \mathbb{R}$, on a

$$\forall x \in \mathbb{E}_n \quad \langle z_{y_1+\alpha y_2}, x\rangle = \langle y_1 + \alpha y_2, u(x)\rangle = \\ \langle y_1, u(x)\rangle + \alpha\langle y_2, u(x)\rangle = \langle z_{y_1} + \alpha z_{y_2}, x\rangle$$

ce qui prouve facilement la linéarité de u^*. ■

REMARQUE 14-3.2 En travaillant dans une base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, on a

$$\forall x \in \mathbb{E}_n \quad u^*(x) = \sum_{i=1}^{n} \langle e_i, u^*(x) \rangle e_i$$

ce qui donne une expression de u^*, compte-tenu de la propriété qui caractérise cet endomorphisme

$$\forall x \in \mathbb{E}_n \quad u^*(x) = \sum_{i=1}^{n} \langle u(e_i), x \rangle e_i$$

14-3.1.2 Propriétés

Les propriétés suivantes découlent facilement de la définition par dualité de l'adjoint :

1. L'application $u \mapsto u^*$ est linéaire et involutive de $\mathcal{L}(\mathbb{E}_n)$ dans lui-même : $(u^*)^* = u$.

   Démonstration : exercice.

2. Pour u et $v \in \mathcal{L}(\mathbb{E}_n)$, on a

   $$(uv)^* = v^* u^*$$

   Démonstration : pour x et $y \in \mathbb{E}_n$, on a

   $$\langle u(v(x)), y \rangle = \langle v(x), u^*(y) \rangle = \langle x, v^*(u^*(y)) \rangle$$

   Cette égalité étant vraie pour tout $x, y \in \mathbb{E}_n$, on a bien $(u \circ v)^* = v^* \circ u^*$.

3. Si $u \in \mathcal{L}(\mathbb{E}_n)$, on a

   $$\ker u^* = (\operatorname{Im} u)^{\perp} \quad \text{et} \quad \operatorname{Im} u^* = (\ker u)^{\perp}$$

   Démonstration : Si $x \in \mathbb{E}_n^*$, le vecteur $u^*(x)$ est nul si et seulement s'il est orthogonal à tout l'espace soit

   $$\forall y \in \mathbb{E}_n \quad \langle u^*(x), y \rangle = \langle x, u(y) \rangle = 0$$

   ce qui traduit exactement $x \in (\operatorname{Im} u)^{\perp}$. Comme $(u^*)^* = u$, on en déduit en remplaçant u par u^*

   $$(\operatorname{Im} u^*)^{\perp} = \ker u \quad \text{et donc} \quad \operatorname{Im} u^* = \left((\operatorname{Im} u^*)^{\perp}\right)^{\perp} = (\ker u)^{\perp}$$

4. Si u est un endomorphisme d'un espace euclidien $\mathbb{E}_n$, u et u^* ont même rang.

   Démonstration : On applique le théorème du rang

   $$\operatorname{rang} u = \dim \operatorname{Im} u = \dim \mathbb{E}_n - \dim (\operatorname{Im} u)^{\perp} = \dim \mathbb{E}_n - \dim \ker u^* = \operatorname{rang} u^*$$

5. Il en résulte que

$$u \in \mathcal{GL}(\mathbb{E}_n) \Leftrightarrow u^* \in \mathcal{GL}(\mathbb{E}_n)$$

et on a alors

$$\left(u^{-1}\right)^* = (u^*)^{-1}$$

Démonstration : u est un automorphisme s'il est de rang n. Il en est alors de même de u^*. On a aussi

$$uu^{-1} = id_{\mathbb{E}_n} \Rightarrow \left(uu^{-1}\right)^* = \left(id_{\mathbb{E}_n}\right)^* = id_{\mathbb{E}_n}$$

ce qui s'écrit aussi $(u^{-1})^* u^* = id_{\mathbb{E}_n}$ et donne bien $(u^{-1})^* = (u^*)^{-1}$.

6. Si $\mathbb{F}$ est un sous-espace de $\mathbb{E}$

$$\mathbb{F} \text{ est } u\text{-stable} \Leftrightarrow \mathbb{F}^{\perp} \text{ est } u^*\text{-stable}$$

Démonstration : Supposons $\mathbb{F}$ stable par u. Si x est un vecteur quelconque de $\mathbb{F}$ et $y \in \mathbb{F}^{\perp}$, on a

$$\langle x, u^*(y)\rangle = \langle u(x), y\rangle = 0$$

puisque $u(x) \in \mathbb{F}$. On a donc bien $u^*(y) \in \mathbb{F}^{\perp}$, ce qui montre que $\mathbb{F}^{\perp}$ est stable par u^*. Comme $(u^*)^* = u$ et $\left(\mathbb{F}^{\perp}\right)^{\perp} = \mathbb{F}$, la réciproque est vraie également.

7. En particulier, la recherche des hyperplans stables par un endomorphisme u se ramène à celle des droites vectorielles stables par u^*, c'est-à-dire à la recherche des vecteurs propres de u^*. Ce résultat est à rapprocher de celui obtenu à la section **5-4.2.5**, valable sur un espace vectoriel de dimension finie sur un corps quelconque. Ce qui suit fait le lien entre ces deux résultats :

8. Si $\mathcal{B}$ est une base **orthonormale**[12] de $\mathbb{E}_n$

$$M(u^*, \mathcal{B}) = {}^tM(u, \mathcal{B})$$

[12]Si $\mathcal{B}$ n'est pas orthonormale et

$$A = (\langle e_i, e_j\rangle)$$

est la matrice du produit scalaire dans $\mathcal{B}$, pour x et $y \in \mathbb{E}_n$ représentés par les colonnes X et Y dans $\mathcal{B}$, la matrice de u dans $\mathcal{B}$ étant M, on a

$$\langle u(x), y\rangle = {}^t(MX)\,AY = {}^tX\,{}^tMAY = {}^tXA\left(A^{-1}\,{}^tMAY\right)$$

quantité qui est aussi égale à

$$\langle x, u^*(y)\rangle = {}^tXA\,(M_1Y)$$

si M_1 est la matrice de u^* dans $\mathcal{B}$. Par unicité de la matrice d'une forme bilinéaire dans une base donnée, on obtient

$$M_1 = A^{-1}\,{}^tMA$$

qui donne bien $M_1 = {}^tM$ lorsque $A = I_n$, c'est-à-dire lorsque la base $\mathcal{B}$ est orthonormale.

Démonstration : Si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ et a_{ij} est le coefficient d'indices (i, j) de la matrice de u^*, on a, puisque la base est orthonormale

$$a_{ij} = \langle e_i, u^*(e_j) \rangle = \langle u(e_i), e_j \rangle = b_{ji}$$

coefficient de la $j^{\text{ième}}$ ligne et $i^{\text{ème}}$ colonne de la matrice $M(u, \mathcal{B})$. ■

9. La propriété qui précède montre que u et u^* ont même déterminant, même polynôme caractéristique et même spectre[13].

REMARQUE 14-3.3 De nombreux résultats obtenus géométriquement peuvent s'obtenir également matriciellement, lorsqu'on sait que, **dans une base orthonormale**, u et u^* ont des matrices transposées l'une de l'autre. Par exemple, si $\mathbb{F}$ est un sous-espace u-stable, dans une base orthonormale de $\mathbb{E}_n$ adaptée à la décomposition

$$\mathbb{E}_n = \mathbb{F} \oplus \mathbb{F}^\perp$$

on a

$$M(u, \mathcal{B}) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

et donc

$$M(u^*, \mathcal{B}) = \begin{pmatrix} {}^tA & 0 \\ {}^tB & {}^tC \end{pmatrix}$$

ce qui montre que $\mathbb{F}^\perp$ est u^*-stable. De plus, si $\mathbb{F}^\perp$ est également u-stable (soit $B = 0$), on a alors $u = u_1 \oplus u_2$ obtenu en recollant deux endomorphismes sur deux supplémentaires **orthogonaux**, on obtient $u^* = u_1^* \oplus u_2^*$.

14-3.1.3 Endomorphismes auto-adjoints

THÉORÈME 14-3.4 (ET DÉFINITION) Si $\mathbb{E}_n$ est un espace euclidien, une application $u : \mathbb{E}_n \to \mathbb{E}_n$ est dite symétrique si et seulement si

$$\forall x, y \in \mathbb{E}_n \quad \langle u(x), y \rangle = \langle x, u(y) \rangle$$

Une telle application est nécessairement linéaire. C'est donc un endomorphisme de $\mathbb{E}_n$ vérifiant

$$u = u^*$$

On dit aussi que u est un endomorphisme auto-adjoint de $\mathbb{E}_n$. Un endomorphisme u est donc auto-adjoint ssi sa matrice <u>dans une base orthonormale</u> est symétrique.

[13]Pour $\lambda \in \mathbb{R}$, on a

$$(u - \lambda id_{\mathbb{E}_n})^* = u^* - \lambda id_{\mathbb{E}_n}$$

et donc

$$\text{rang}(u^* - \lambda id_{\mathbb{E}_n}) = \text{rang}(u - \lambda id_{\mathbb{E}_n})$$

Les sous-espaces propres de u et u^* pour une valeur propre λ donnée ont donc même dimension. Rappelons qu'à un vecteur propre de u^* correspond un hyperplan stable par u.

Démonstration : Si $x_1, x_2 \in \mathbb{E}_n$ et $\alpha \in \mathbb{R}$, montrer que

$$u(x_1 + \alpha x_2) = u(x_1) + \alpha u(x_2)$$

revient à prouver l'égalité des formes linéaires associées

$$\langle u(x_1 + \alpha x_2), \bullet \rangle = \langle u(x_1) + \alpha u(x_2), \bullet \rangle$$

Or, pour $y \in \mathbb{E}_n$ quelconque

$$\langle u(x_1 + \alpha x_2), y \rangle = \langle x_1 + \alpha x_2, u(y) \rangle = \\ \langle x_1, u(y) \rangle + \alpha \langle x_2, u(y) \rangle = \langle u(x_1), y \rangle + \alpha \langle u(x_2), y \rangle$$

ce qui prouve le résultat. ■

De même, une application $u : \mathbb{E}_n \to \mathbb{E}_n$ est antisymétrique si et seulement si

$$\forall x, y \in \mathbb{E}_n \quad \langle u(x), y \rangle = -\langle x, u(y) \rangle$$

Cette propriété entraîne également la linéarité de u (exercice) et se traduit donc par l'égalité $u = -u^*$.

PROPOSITION 14-3.5 L'ensemble $\mathcal{S}(\mathbb{E}_n)$ des endomorphismes symétriques et l'ensemble $\mathcal{A}(\mathbb{E}_n)$ des endomorphismes antisymétriques de $\mathbb{E}_n$ sont deux sous-espaces supplémentaires de $\mathcal{L}(\mathbb{E}_n)$, isomorphes respectivement aux espaces de matrices $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$. Ils sont de dimensions respectives $\dfrac{n(n+1)}{2}$ et $\dfrac{n(n-1)}{2}$.

Démonstration : L'application $\phi : \mathcal{L}(\mathbb{E}_n) \to \mathcal{L}(\mathbb{E}_n)$ définie par $\phi(u) = u^*$ est une involution linéaire. On a donc

$$\mathcal{L}(\mathbb{E}_n) = \ker\left(\phi - id_{\mathcal{L}(\mathbb{E}_n)}\right) \oplus \ker\left(\phi + id_{\mathcal{L}(\mathbb{E}_n)}\right) = \mathcal{S}(\mathbb{E}_n) \oplus \mathcal{A}(\mathbb{E}_n)$$

la décomposition d'un morphisme u dans cette somme directe est simplement

$$u = \frac{1}{2}(u + u^*) + \frac{1}{2}(u - u^*)$$

Le choix d'une base $\mathcal{B}$ orthonormale de l'espace fournit un isomorphisme

$$u \mapsto M(u, \mathcal{B})$$

de $\mathcal{S}(\mathbb{E}_n)$ vers $\mathcal{S}_n(\mathbb{R})$ et de $\mathcal{A}(\mathbb{E}_n)$ vers $\mathcal{A}_n(\mathbb{R})$. ■

Un exemple fondamental d'endomorphisme symétrique est fourni par les projecteurs orthogonaux :

PROPOSITION 14-3.6 Un projecteur $p \in \mathcal{L}(\mathbb{E}_n)$ est auto-adjoint si et seulement si c'est un projecteur orthogonal : il existe $\mathbb{F}$ sous-espace vectoriel de $\mathbb{E}$ tel que p soit le projecteur orthogonal sur $\mathbb{F}$, soit

$$\operatorname{Im} p = \mathbb{F} \quad \text{et} \quad \ker p = \mathbb{F}^{\perp}$$

Démonstration : Si p est auto-adjoint, on a, d'après la propriété 3 de la section **14-3.1.2**

$$\ker p = \ker p^* = (\operatorname{Im} p)^\perp$$

et p est donc un projecteur orthogonal. Réciproquement, si p est projecteur orthogonal sur $\mathbb{F}$ et si deux vecteurs x et y se décomposent en $x = x' + x''$ et $y = y' + y''$ dans la somme directe $\mathbb{F} \oplus \mathbb{F}^\perp$, on a

$$\langle p(x), y\rangle = \langle x', y' + y''\rangle = \langle x', y'\rangle = \langle x, p(y)\rangle$$

et p est bien symétrique. ■

Comme $\mathcal{S}(\mathbb{E}_n)$ est un espace vectoriel, on en déduit que, si $(p_i)_{1 \leqslant i \leqslant m}$ est la famille de projecteurs associée à une décomposition **orthogonale** de l'espace

$$\mathbb{E}_n = \mathbb{F}_1 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} \mathbb{F}_m$$

tout endomorphisme s'écrivant

$$u = \sum_{i=1}^{m} \lambda_i p_i$$

avec $(\lambda_i)_{1 \leqslant i \leqslant m} \in \mathbb{R}^m$ est auto-adjoint. Le théorème de réduction des endomorphismes symétriques (section **14-3.3**) nous montrera que tous les endomorphismes symétriques peuvent s'écrire de cette façon.

En particulier, si $\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^\perp$ et p et q sont les deux projecteurs orthogonaux asssociés à cette décomposition de l'espace, la symétrie orthogonale par rapport à $\mathbb{F}$

$$s = p - q = 2p - id_{\mathbb{E}}$$

est un automorphisme (involutif) auto-adjoint. Cette transformation linéaire de l'espace conserve la norme, puisque si $x = x_1 + x_2$ est la décomposition d'un vecteur quelconque de l'espace dans $\mathbb{F} \oplus \mathbb{F}^\perp$, on a $s(x) = x_1 - x_2$, et par la relation de Pythagore

$$\|s(x)\|^2 = \|x_1\|^2 + \|x_2\|^2 = \|x\|^2$$

Les symétries orthogonales d'un espace euclidien sont des exemples d'automorphismes orthogonaux, que nous étudions dans la section qui suit.

COROLLAIRE 14-3.7 Dans une base orthonormale d'un espace euclidien, la matrice d'un projecteur orthogonal ou d'une symétrie orthogonale est symétrique.

EXERCICE 14-3.8 Montrer qu'un projecteur p d'un espace euclidien $\mathbb{E}_n$ est orthogonal si et seulement si

$$\forall x \in \mathbb{E}_n \quad \|p(x)\| \leqslant \|x\|$$

14-3.2 Automorphismes orthogonaux

14-3.2.1 Définition

DÉFINITION 14-3.9 *On appelle automorphisme orthogonal d'un espace euclidien $\mathbb{E}_n$ toute application $u : \mathbb{E}_n \to \mathbb{E}_n$ conservant le produit scalaire*

$$\forall x, y \in \mathbb{E}_n \quad \langle u(x), u(y) \rangle = \langle x, y \rangle$$

La terminologie utilisée s'explique par le fait qu'une application qui conserve le produit scalaire est nécessairement linéaire et bijective : si u possède cette propriété, pour $\alpha \in \mathbb{R}$ et $x, y \in \mathbb{E}_n$ on a

$$\begin{aligned} &\|u(x+\alpha y) - u(x) - \alpha u(y)\|^2 = \\ &\qquad \|u(x+\alpha y)\|^2 + \|u(x)\|^2 + \alpha^2 \|u(y)\|^2 + 2\alpha \langle u(x), u(y) \rangle \\ &\qquad\qquad - 2\langle u(x+\alpha y), u(x) \rangle - 2\alpha \langle u(x+\alpha y), u(y) \rangle \end{aligned}$$

ce qui donne, par conservation du produit scalaire (donc aussi du carré scalaire)

$$\|u(x+\alpha y) - u(x) - \alpha u(y)\|^2 = \|(x+\alpha y) - x - \alpha y\|^2 = 0$$

et prouve la linéarité de u. De plus, si $u(x) = 0_{\mathbb{E}_n}$, on a

$$\|u(x)\|^2 = \|x\|^2 = 0 \Rightarrow x = 0_{\mathbb{E}_n}$$

d'où l'injectivité de u. Comme l'espace $\mathbb{E}_n$ est supposé de dimension finie[14], $u \in \mathcal{GL}(\mathbb{E}_n)$. Nous noterons $\mathcal{O}(\mathbb{E}_n)$ l'ensemble des automorphismes orthogonaux de l'espace euclidien $\mathbb{E}_n$.

14-3.2.2 Caractérisation. Propriétés

Pour prouver qu'un endomorphisme d'un espace euclidien est un automorphisme orthogonal, nous disposons de nombreuses caractérisations :

THÉORÈME 14-3.10 Soit u un endomorphisme d'un espace euclidien $\mathbb{E}_n$. Les propriétés suivantes sont équivalentes :

i) u conserve la norme

$$\forall x \in \mathbb{E}_n \quad \|u(x)\| = \|x\|$$

ii) u conserve le produit scalaire (donc $u \in \mathcal{O}(\mathbb{E}_n)$).

[14] En dimension infinie, u n'est pas nécessairement surjectif. Considérer par exemple l'application de $l^2(\mathbb{N})$ dans lui-même définie par

$$(u_0, u_1, \ldots, u_n, \ldots) \mapsto (0, u_0, u_1, \ldots, u_n, \ldots)$$

si $l^2(\mathbb{N})$ est muni du produit scalaire usuel

$$\langle (u_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}} \rangle = \sum_{n=0}^{+\infty} u_n v_n$$

iii) u est un automorphisme de $\mathbb{E}_n$ et $u^{-1} = u^*$

$$uu^* = u^*u = id_{\mathbb{E}_n}$$

iv) La matrice M de u dans une base orthonormale de $\mathbb{E}_n$ vérifie

$$^tMM = I_n$$

v) Il existe une base orthonormale transformée par u en base orthonormale.

vi) L'image de toute base orthonormale de $\mathbb{E}_n$ est une base orthonormale.

Démonstration : Si u est linéaire et conserve la norme, la conservation du produit scalaire s'obtient par "polarisation" :

$$\forall\, x, y \in \mathbb{E}_n \quad \langle u(x), u(y)\rangle = \frac{1}{4}\left(\|u(x) + u(y)\|^2 - \|u(x) - u(y)\|^2\right)$$

La linéarité de u donne $u(x \pm y) = u(x) \pm u(y)$, et on obtient

$$\begin{aligned}\langle u(x), u(y)\rangle &= \frac{1}{4}\left(\|u(x+y)\|^2 - \|u(x-y)\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) = \langle x, y\rangle\end{aligned}$$

On a donc bien $i) \Rightarrow ii)$.

Si $u \in \mathcal{O}(\mathbb{E}_n)$, on sait que u est un automorphisme. Comme u conserve le produit scalaire, on a, pour x et $y \in \mathbb{E}_n$

$$\left\langle x, u^{-1}(y)\right\rangle = \left\langle u(x), u\left(u^{-1}(y)\right)\right\rangle = \langle u(x), y\rangle$$

ce qui montre que $u^* = u^{-1}$. On a prouvé $ii) \Rightarrow iii)$.

Si M est matrice de u dans une base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, la matrice de u^* dans cette même base est ^tM. La propriété $iv)$ est donc la simple traduction matricielle de $iii)$.

Si M, matrice de u dans une base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, vérifie $^tMM = I_n$, le coefficient d'indices (i, j) de la matrice produit ^tMM est produit de la $i^{\text{ème}}$ ligne de ^tM par la $j^{\text{ème}}$ colonne de M. C'est donc le produit scalaire (canonique dans $\mathbb{R}^n$) de la $i^{\text{ème}}$ colonne de M et de la $j^{\text{ème}}$ colonne de M. Comme $\mathcal{B}$ est supposé orthonormale, c'est

$$\sum_{k=1}^{n} \langle u(e_i), e_k\rangle \langle u(e_j), e_k\rangle = \langle u(e_i), u(e_j)\rangle = \delta_i^j$$

puisque par hypothèse la matrice produit est I_n. La famille $(u(e_i))_{1 \leqslant i \leqslant n}$ est donc une famille orthonormale de cardinal n, c'est une base orthonormale de $\mathbb{E}_n$. On a bien $iv) \Rightarrow v)$.

Enfin, si u est un endomorphisme de $\mathbb{E}_n$ qui transforme une base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ en une base orthonormale, on a pour $x, y \in \mathbb{E}_n$ décomposés dans la base $\mathcal{B}$

$$x = \sum_{i=1}^{n} x_i e_i \Rightarrow \|x\|^2 = \sum_{i=1}^{n} x_i^2 = \left\|\sum_{i=1}^{n} x_i u(e_i)\right\|^2 = \|u(x)\|^2$$

$$y = \sum_{i=1}^{n} y_i e_i \Longrightarrow \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$
$$= \left\langle \sum_{i=1}^{n} x_i u(e_i), \sum_{i=1}^{n} y_i u(e_i) \right\rangle = \langle u(x), u(y) \rangle$$

Donc u conserve la norme et le produit scalaire et transforme toute famille orthonormale de cardinal n en une famille de même type : on a donc bien $v) \Rightarrow vi)$ et *a fortiori* $vi) \Rightarrow i)$. ■

REMARQUE 14-3.11 Si $u : \mathbb{E}_n \to \mathbb{E}_n$ est une application quelconque qui conserve la norme, u n'est en général pas linéaire : considérer par exemple un vecteur $e \in \mathbb{E}_n$ vérifiant $\|e\| = 1$ et l'application u définie par

$$\forall x \in \mathbb{E}_n \quad u(x) = \|x\| e$$

Par contre, si on impose à u de conserver la distance euclidienne

$$\forall x, y \in \mathbb{E}_n \quad \|u(x) - u(y)\| = \|x - y\|$$

alors l'application $x \mapsto v(x) = u(x) - u(0_{\mathbb{E}_n})$ est une application linéaire qui conserve la norme, donc un automorphisme orthogonal, et u est donc une application affine associée à la transformation vectorielle v. En effet, on a par construction $v(0_{\mathbb{E}_n}) = 0_{\mathbb{E}_n}$ et, pour x et $y \in \mathbb{E}_n$

$$\|v(x) - v(y)\|^2 = \|x - y\|^2$$

ce qui entraîne en particulier que v conserve la norme (faire $y = 0_{\mathbb{E}_n}$). En développant l'égalité précédente, on obtient également

$$\langle v(x), v(y) \rangle = \langle x, y \rangle$$

propriété qui entraîne $v \in \mathcal{O}(\mathbb{E}_n)$. **Toute application d'un espace euclidien dans lui-même conservant les distances est une isométrie affine.**

La conservation de la norme et du produit scalaire ont des conséquences importantes, qu'il faut avoir à l'esprit lorsqu'on cherche à décrire un automorphisme orthogonal :

- Si $\mathbb{F}$ est un sous-espace vectoriel stable par un automorphisme orthogonal u, son supplémentaire orthogonal est u-stable également, et les endomorphismes induits sont eux aussi des automorphismes orthogonaux

  $$u|_{\mathbb{F}} \in \mathcal{O}(\mathbb{F}) \quad \text{et} \quad u|_{\mathbb{F}^\perp} \in \mathcal{O}(\mathbb{F}^\perp)$$

  En effet, on sait que si $\mathbb{F}$ est u-stable son supplémentaire orthogonal est u^*-stable. Comme $u^* = u^{-1}$, on a

  $$u^{-1}(\mathbb{F}^\perp) \subset \mathbb{F}^\perp$$

  ce qui donne $\mathbb{F}^\perp \subset u(\mathbb{F}^\perp)$. On a en fait une égalité pour des raisons de dimension. Comme u conserve la norme, il en est évidemment de même pour $u|_{\mathbb{F}}$ et $u|_{\mathbb{F}^\perp}$.

- Réciproquement, si $\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^{\perp}$ est une décomposition orthogonale de l'espace, le recollement linéaire d'un automorphisme orthogonal $u_1 \in \mathcal{O}(\mathbb{F})$ et d'un automorphisme orthogonal $u_2 \in \mathcal{O}\left(\mathbb{F}^{\perp}\right)$ donne un élément de $\mathcal{O}(\mathbb{E}_n)$: si $u = u_1 \oplus u_2$ et si $x \in \mathbb{E}_n$ est décomposé en $x_1 + x_2$ dans $\mathbb{F} \oplus \mathbb{F}^{\perp}$, on a par la relation de Pythagore

$$\begin{aligned} \|u(x)\|^2 &= \|u_1(x_1) + u_2(x_2)\|^2 \\ &= \|u_1(x_1)\|^2 + \|u_2(x_2)\|^2 = \|x_1\|^2 + \|x_2\|^2 = \|x\|^2 \end{aligned}$$

ce qui donne bien $u \in \mathcal{O}(\mathbb{E}_n)$.

- Si $u \in \mathcal{O}(\mathbb{E}_n)$, les seules valeurs propres possibles de u sont ± 1

$$\operatorname{Sp} u \subset \{\pm 1\}$$

et on a une décomposition orthogonale de l'espace en sous-espaces u-stables

$$\mathbb{E}_n = \ker(u - id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \ker(u + id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \mathbb{S}$$

avec $u|_{\mathbb{S}} \in \mathcal{O}(\mathbb{S})$ n'ayant aucune valeur propre. En effet, si $x \in \mathbb{E}_n$ est vecteur propre de u pour la valeur propre λ, on a $u(x) = \lambda x$ et $\|u(x)\| = |\lambda| \|x\| = \|x\| \Rightarrow \lambda = \pm 1$. De plus, si $u(x) = x$ et $u(y) = -y$, la conservation du produit scalaire donne

$$\langle x, y\rangle = -\langle x, y\rangle \Rightarrow \langle x, y\rangle = 0$$

Enfin, $\ker(u - id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \ker(u + id_{\mathbb{E}_n})$ étant un sous-espace u-stable, son supplémentaire orthogonal $\mathbb{S}$ est aussi u-stable.

EXERCICE 14-3.12 Si A est la matrice du produit scalaire dans une base $\mathcal{B}$ quelconque, montrer qu'un endomorphisme u de $\mathbb{E}_n$ est un automorphisme orthogonal ssi sa matrice $M = M(u, \mathcal{B})$ vérifie

$${}^tMAM = A$$

14-3.2.3 Groupe orthogonal

PROPOSITION 14-3.13 L'ensemble $\mathcal{O}(\mathbb{E}_n)$ des automorphismes orthogonaux d'un espace euclidien muni de la composition est un sous-groupe du groupe linéaire $(\mathcal{GL}(\mathbb{E}_n), \circ)$, appelé groupe orthogonal de l'espace $(\mathbb{E}_n, \langle\ ,\ \rangle)$.

Démonstration : Exercice.

PROPOSITION 14-3.14 L'application $\det : (\mathcal{O}(\mathbb{E}_n), \circ) \to (\{\pm 1\}, \times)$ **est un morphisme surjectif de groupes. Son noyau**

$$\mathcal{O}^+(\mathbb{E}_n) = \{u \in \mathcal{O}(\mathbb{E}_n) \mid \det u = 1\}$$

est un sous-groupe de $\mathcal{O}(\mathbb{E}_n)$, appelé groupe spécial orthogonal de l'espace euclidien $\mathbb{E}_n$.

Démonstration : Si $\mathcal{B}$ est une base orthonormale de $\mathbb{E}_n$ et u est un automorphisme orthogonal, la matrice M de u dans $\mathcal{B}$ vérifie la relation ${}^tMM = I_n$, et donc

$$\det\left({}^tMM\right) = [\det M]^2 = 1$$

ce qui entraîne $\det u = \pm 1$. Si x est un vecteur non nul de $\mathbb{E}_n$, la symétrie orthogonale par rapport à l'hyperplan $x^\perp$ est dans $\mathcal{O}(\mathbb{E}_n)$, et de déterminant égal à -1, ce qui prouve la surjectivité de l'application déterminant $(\mathcal{O}(\mathbb{E}_n), \circ) \to (\{\pm 1\}, \times)$. Le noyau de ce morphisme est un sous-groupe de $(\mathcal{O}(\mathbb{E}_n), \circ)$. ■

Les éléments de $\mathcal{O}^+(\mathbb{E}_n)$ sont appelés automorphismes orthogonaux directs[15] de $\mathbb{E}_n$. L'ensemble

$$\mathcal{O}^-(\mathbb{E}_n) = \{u \in \mathcal{O}(\mathbb{E}_n) \mid \det u = -1\}$$

est l'ensemble des automorphismes orthogonaux indirects de $\mathbb{E}_n$. Il n'est pas stable par composition, puisque le produit de deux éléments de $\mathcal{O}^-(\mathbb{E}_n)$ est dans $\mathcal{O}^+(\mathbb{E}_n)$. Les réflexions sont un exemple important de transformations orthogonales indirectes :

DÉFINITION 14-3.15 *On appelle réflexion d'un espace euclidien toute symétrie orthogonale par rapport à un hyperplan.*

Si cet hyperplan est $\mathbb{H} = x^\perp$ avec x vecteur non nul de $\mathbb{E}_n$, dans une base orthonormale adaptée à la décomposition

$$\mathbb{E}_n = \operatorname{vect} x \oplus \mathbb{H}$$

la matrice de cette réflexion est simplement $\operatorname{Diag}(-1, 1, \ldots, 1)$.

L'exercice qui suit montre que les réflexions engendrent le groupe $\mathcal{O}(\mathbb{E}_n)$, c'est-à-dire que tout automorphisme orthogonal d'un espace euclidien peut s'écrire comme produit de réflexions. La démonstration utilise le résultat intermédiaire :

LEMME 14-3.16 Si a et b sont deux vecteurs distincts d'un espace euclidien vérifiant

$$\|a\| = \|b\|$$

il existe une unique réflexion de l'espace échangeant a et b. Si $x \in \mathbb{E}_n$, son image par cette réflexion est donnée par

$$s(x) = x - 2\,(b-a, x)\,\frac{b-a}{\|b-a\|^2}$$

Démonstration : Si s est une telle réflexion, on a

$$s(b-a) = s(b) - s(a) = a - b$$

[15] On dit aussi "rotations".

$b - a \neq 0_{\mathbb{E}_n}$ engendre donc la droite vectorielle propre de s pour la valeur propre -1, et s est donc la réflexion par rapport à l'hyperplan orthogonal $\mathbb{H} = (b-a)^{\perp}$. Réciproquement, si s est cette réflexion, on a $s(b-a) = a - b$. Mais $\|a\| = \|b\|$, et on a donc

$$\langle b - a, b + a \rangle = 0$$

donc $s(b+a) = b + a$, ce qui donne finalement $s(a) = b$ et $s(b) = a$. On a de plus

$$s = id_{\mathbb{E}_n} - 2p$$

où p est le projecteur orthogonal sur $\text{vect}(b-a)$, ce qui donne facilement l'expression de $s(x)$. ■

EXERCICE 14-3.17 En raisonnant par récurrence sur la dimension de l'espace et en utilisant le résultat précédent, montrer que toute transformation orthogonale ($\neq id_{\mathbb{E}_n}$) d'un espace euclidien peut s'écrire comme produit de réflexions. Montrer que le nombre de réflexions intervenant dans la décomposition d'un automorphisme orthogonal peut être pris $\leqslant n$ (dimension de l'espace).

EXERCICE 14-3.18 Si $\mathbb{R}.a$ et $\mathbb{R}.b$ sont deux droites vectorielles distinctes d'un espace euclidien, montrer qu'il existe exactement deux réflexions qui échangent ces droites.

14-3.2.4 Matrices orthogonales

DÉFINITION 14-3.19 *Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si*

$$^tMM = I_n$$

On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales de type (n,n). Si $\mathbb{R}^n$ est muni de son produit scalaire canonique, dire que M est orthogonale revient exactement à dire que l'endomorphisme canoniquement associé est dans $\mathcal{O}(\mathbb{R}^n)$. Il en résulte immédiatement :

PROPOSITION 14-3.20 L'ensemble $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe du groupe linéaire $(\mathcal{GL}_n(\mathbb{R}), \times)$. Il est isomorphe au groupe orthogonal de l'espace $\mathbb{R}^n$ muni de sa structure euclidienne canonique. Le groupe $(\mathcal{O}_n(\mathbb{R}), \times)$ est appelé groupe orthogonal d'ordre n.

Plus généralement, si $\mathcal{B}$ est une base orthonormale d'un espace $(\mathbb{E}_n, \langle\, , \rangle)$ euclidien de dimension n, le choix de cette base donne un isomorphisme de groupes

$$\begin{aligned} \mathcal{O}(\mathbb{E}_n) &\to \mathcal{O}_n(\mathbb{R}) \\ u &\mapsto M(u, \mathcal{B}) \end{aligned}$$

Si $M \in \mathcal{O}_n(\mathbb{R})$, on a

$$\det M = \pm 1$$

et l'ensemble des matrices orthogonales directes

$$\mathcal{O}_n^+(\mathbb{R}) = \{M \in \mathcal{O}_n(\mathbb{R}) \mid \det M = 1\}$$

est un sous-groupe de $(\mathcal{O}_n(\mathbb{R}), \times)$, appelé groupe spécial orthogonal d'ordre n. Le choix d'une base orthonormale d'un espace euclidien $\mathbb{E}_n$ crée un isomorphisme $u \mapsto M(u, \mathcal{B})$ de $(\mathcal{O}^+(\mathbb{E}_n), \circ)$ vers $(\mathcal{O}_n^+(\mathbb{R}), \times)$.

On peut interpréter simplement l'égalité ${}^tMM = I_n$ caractérisant les matrices carrées orthogonales : comme nous l'avons vu à la section **14-3.2.2**, si

$$M = (C_1, \ldots, C_n)$$

est considérée comme matrice de vecteurs colonnes, le coefficient d'indice (i, j) de la matrice ${}^tMM$ est simplement

$$a_{ij} = {}^tC_iC_j = (C_i|C_j)$$

produit scalaire canonique de ces deux colonnes. On a donc

PROPOSITION 14-3.21 Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ses colonnes forment une base orthonormale de $\mathbb{R}^n$ muni de sa structure euclidienne canonique. Cette propriété est vérifiée aussi par les vecteurs lignes de M, puisque ${}^tMM = I_n$ équivaut à $M{}^tM = I_n$.

EXERCICE 14-3.22 $\mathcal{M}_n(\mathbb{R})$ étant muni d'une norme quelconque, montrer que $\mathcal{O}_n(\mathbb{R})$ est une partie compacte de $\mathcal{M}_n(\mathbb{R})$. Est-ce une partie connexe ?

Pour le moment, nous avons considéré les matrices orthogonales comme matrices, dans une base orthonormale, d'automorphismes orthogonaux. On peut aussi les considérer comme matrices de passage entre bases orthonormales :

PROPOSITION 14-3.23 Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base orthonormale d'un espace euclidien $(\mathbb{E}_n, \langle\, ,\, \rangle)$ et $\mathcal{B}' = (e'_i)_{1\leqslant i\leqslant n}$ une autre base, $P = \mathcal{M}_{\mathcal{B}}(\mathcal{B}')$ étant la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. On a

$$\mathcal{B}' \text{ orthonormale} \Leftrightarrow P \in \mathcal{O}_n(\mathbb{R})$$

Démonstration : Les colonnes de P représentent les coordonnées des vecteurs de $\mathcal{B}'$ dans $\mathcal{B}$. Comme $\mathcal{B}$ est orthonormale, on a

$$\langle e'_i, e'_j \rangle = (C_i|C_j)$$

et donc l'égalité ${}^tMM = I_n$ traduit exactement le fait que $\mathcal{B}'$ soit orthonormale. ■

Les formules de changement de bases orthonormales dans un espace euclidien sont donc simples : si $x \in \mathbb{E}_n$ est représenté dans $\mathcal{B}$ par la colonne X et dans $\mathcal{B}'$ par X', on a

$$X = PX' \Leftrightarrow X' = {}^tPX$$

REMARQUE 14-3.24 Si M est une matrice orthogonale, on a

$$M^{-1} = {}^tM = \frac{1}{\det M}{}^t(\operatorname{com} M)$$

et donc $M = \pm \operatorname{com} M$, le signe étant $+$ lorsque M est directe. Pour vérifier qu'une matrice de $\mathcal{M}_n(\mathbb{R})$ est orthogonale directe, il suffit donc de vérifier que ses colonnes ou ses lignes forment une base orthonormale de $\mathbb{R}^n$ euclidien canonique et qu'un de ses coefficients non nuls est égal à son cofacteur (ce qui amène au calcul d'un déterminant d'ordre $n-1$ et évite de calculer un déterminant d'ordre n).

14-3.3 Réduction des endomorphismes symétriques

14-3.3.1 Diagonalisation des endomorphismes auto-adjoints

Si u est un endomorphisme symétrique d'un espace euclidien $(\mathbb{E}_n, \langle\, ,\, \rangle)$, on a, pour tout $\lambda \in \mathbb{R}$

$$(u - \lambda id_{\mathbb{E}_n})^* = u - \lambda id_{\mathbb{E}_n}$$

ce qui donne en particulier

$$\ker(u - \lambda id_{\mathbb{E}_n}) = \operatorname{Im}(u - \lambda id_{\mathbb{E}_n})^{\perp}$$

et par conséquent, les sous-espaces $\ker(u - \lambda id_{\mathbb{E}_n})$ et $\operatorname{Im}(u - \lambda id_{\mathbb{E}_n})$ sont supplémentaires orthogonaux[16] :

$$\forall \lambda \in \mathbb{R} \quad \mathbb{E}_n = \operatorname{Im}(u - \lambda id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \ker(u - \lambda id_{\mathbb{E}_n})$$

Il en résulte en particulier que tout sous-espace propre d'un endomorphisme symétrique possède un supplémentaire stable, résultat qui est à la base du théorème fondamental qui suit.

LEMME 14-3.25 Si u est un endomorphisme symétrique de $(\mathbb{E}_n, \langle\, ,\, \rangle)$, les sous-espaces propres associés à des valeurs propres distinctes sont orthogonaux.

Démonstration : Si $\lambda_1 \neq \lambda_2$ sont deux valeurs propres de f, on a pour des vecteurs $x_1 \in \ker(u - \lambda_1 id_{\mathbb{E}_n})$ et $x_2 \in \ker(u - \lambda_2 id_{\mathbb{E}_n})$

$$\langle u(x_1), x_2\rangle = \lambda_1 \langle x_1, x_2\rangle = \langle x_1, u(x_2)\rangle = \lambda_2 \langle x_1, x_2\rangle$$

ce qui donne bien $\langle x_1, x_2\rangle = 0$. On aurait pu aussi utiliser ce qui précède puisqu'il est clair que, pour un endomorphisme u quelconque et deux scalaires λ_1 et λ_2 distincts,

$$\ker(u - \lambda_1 id_{\mathbb{E}_n}) \subset \operatorname{Im}(u - \lambda_2 id_{\mathbb{E}_n})$$

et, si u est auto-adjoint, $\operatorname{Im}(u - \lambda_2 id_{\mathbb{E}_n}) = \ker(u - \lambda_2 id_{\mathbb{E}_n})^{\perp}$. ■

[16] En particulier, le noyau et l'image d'un endomorphisme autoadjoint sont toujours supplémentaires orthogonaux.

Il résulte de ce lemme que, si $u \in \mathcal{S}(\mathbb{E}_n)$, le sous-espace[17]

$$\mathbb{F} = \overset{\perp}{\bigoplus_{\lambda \in \mathrm{spec}\, u}} \ker(u - \lambda id_{\mathbb{E}_n})$$

est un sous-espace u-stable, dont le supplémentaire orthogonal $\mathbb{F}^{\perp}$ est u^*-stable, donc u-stable. L'endomorphisme induit $v = u|_{\mathbb{F}^{\perp}}$ est dans $\mathcal{S}(\mathbb{F}^{\perp})$, et est manifestement de spectre vide. Le lemme qui suit nous montrera que, nécessairement, $\mathbb{F}^{\perp}$ est réduit au vecteur nul. Il en résultera que l'espace $\mathbb{E}_n$ est somme directe orthogonale des sous-espaces propres de u, et donc que u est diagonalisable.

LEMME 14-3.26 Si u est un endomorphisme auto-adjoint d'un espace euclidien de dimension $n > 0$, le spectre de u n'est pas vide.

Démonstration : Nous donnerons une démonstration matricielle de ce résultat, démonstration qui prendra tout son sens lorsque le chapitre sur les espaces hermitiens aura été étudié. L'exercice qui suit donne une démonstration un peu plus géométrique. Voir aussi l'exercice **18-3**.42. La matrice de u dans une base orthonormale de l'espace est une matrice $A \in \mathcal{S}_n(\mathbb{R})$. Le polynôme caractéristique de A est dans $\mathbb{R}[X]$, et nous montrons qu'il est scindé dans $\mathbb{R}[X]$. En effet, si $\lambda \in \mathbb{C}$ est racine de ce polynôme, on peut trouver un vecteur colonne $X \in \mathbb{C}^n$ non nul avec $AX = \lambda X$. Comme A est à coefficients réels, on a par conjugaison

$$A\overline{X} = \overline{\lambda X}$$

ce qui va donner

$${}^tXA\overline{X} = {}^t(AX)\overline{X} = \lambda{}^tX\overline{X} = {}^tX\left(A\overline{X}\right) = \overline{\lambda}{}^tX\overline{X}$$

On obtient finalement

$$(\lambda - \overline{\lambda})\left(\sum_{i=1}^{n} |x_i|^2\right) = 0 \quad \text{si} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Comme X est non nul, on obtient $\left(\lambda - \overline{\lambda}\right) = 0$, soit $\lambda \in \mathbb{R}$. ■

EXERCICE 14-3.27 Une autre démonstration de ce résultat : montrer qu'en dimension 2, tout endomorphisme symétrique possède au moins une valeur propre (difficile d'échapper au calcul matriciel). Si $u \in \mathcal{S}(\mathbb{E}_n)$ avec $n > 2$ ne possède aucune valeur propre, montrer que son polynôme minimal a tous ses diviseurs irréductibles (dans $\mathbb{R}[X]$) de degré 2. Si P est un tel diviseur, montrer que $\ker P(u) \neq \{0_{\mathbb{E}_n}\}$, et que si on choisit un vecteur non nul dans ce sous-espace, le plan $\mathrm{vect}(x, u(x))$ est u-stable. Conclure.

[17] Avec, par convention $\mathbb{F} = \{0_{\mathbb{E}_n}\}$ si le spectre de u était vide !

THÉORÈME 14-3.28 Tout endomorphisme symétrique u d'un espace euclidien $\mathbb{E}_n$ est diagonalisable. Plus précisément, l'espace $\mathbb{E}_n$ est somme directe orthogonale des différents sous-espaces propres de u. La famille obtenue par réunion de bases orthonormales de ces différents sous-espaces est alors clairement une base orthonormale de $\mathbb{E}_n$ qui diagonalise u.

COROLLAIRE 14-3.29 Un endomorphisme d'un espace euclidien est auto-adjoint si et seulement s'il est diagonalisable dans une base orthonormale.

> *Démonstration* : Cette condition est nécessaire d'après ce qui précède. Si elle est vérifiée, il existe une base orthonormale où la matrice de l'endomorphisme considéré est diagonale, donc symétrique. L'endomorphisme est lui-même symétrique. ■

On retrouve bien sûr le fait qu'un projecteur orthogonal est auto-adjoint. Plus précisément, le théorème qui précède peut s'écrire, en considérant la famille de projecteurs spectraux d'un endomorphisme symétrique :

COROLLAIRE 14-3.30 Un endomorphisme d'un espace euclidien est auto-adjoint si et seulement s'il est combinaison linéaire d'une famille de projecteurs orthogonaux.

EXERCICE 14-3.31 Montrer que, si p_1 et p_2 sont deux projecteurs orthogonaux d'un espace euclidien, leur composé $p_1 \circ p_2$ est diagonalisable (on pourra utiliser des vecteurs propres de $p_1 + p_2$).

14-3.3.2 Traduction matricielle

Dans $\mathbb{R}^n$ muni de sa structure euclidienne canonique, l'endomorphisme canoniquement associé à une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique si et seulement si $A = {}^tA$. La matrice de passage de la base canonique à une autre base est orthogonale si et seulement si cette autre base est orthonormale. Le théorème **14-3**.28 s'écrit alors

THÉORÈME 14-3.32 Toute matrice A de $\mathcal{S}_n(\mathbb{R})$ est diagonalisable. Plus précisément, si $A \in \mathcal{S}_n(\mathbb{R})$, on peut trouver une matrice de passage O orthogonale telle que

$$D = O^{-1}AO = {}^tOAO$$

soit une matrice diagonale[18].

EXERCICE 14-3.33 On peut imposer à O d'être dans $\mathcal{O}_n^+(\mathbb{R})$.

[18] Cette propriété caractérise les matrices réelles symétriques, puisque si O existe, la matrice A est semblable, mais aussi congruente (car $O^{-1} = {}^tO$) à une matrice diagonale, et est donc symétrique.

REMARQUE 14-3.34 Ce théorème est un résultat **réel**. Il n'est pas valable sur un corps quelconque, comme le montre l'exemple

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$$

A est une matrice symétrique non diagonalisable. Nous verrons dans le prochain chapitre comment remplacer l'hypothèse $A = {}^tA$ pour avoir un théorème analogue sur $\mathbb{C}$.

EXERCICE 14-3.35 Donner une matrice de passage orthogonale diagonalisant la matrice réelle

$$A = \begin{pmatrix} a & 1 & \cdots & \cdots & 1 \\ 1 & \ddots & 1 & 1 & \vdots \\ \vdots & 1 & a & 1 & \vdots \\ \vdots & 1 & 1 & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & a \end{pmatrix}$$

14-3.4 Formes quadratiques sur un espace euclidien

Dans une base **orthonormale** d'un espace euclidien $(\mathbb{E}_n, \langle\ ,\ \rangle)$, un endomorphisme symétrique u possède une matrice symétrique. Nous allons voir dans cette section que cette matrice peut également être interprétée comme matrice, dans cette base, d'une forme quadratique Φ sur $\mathbb{E}_n$ et que la correspondance $u \mapsto \Phi$ est "intrinsèque", c'est-à-dire indépendante de la base orthonormale choisie.

14-3.4.1 Endomorphisme auto-adjoint associé à une forme quadratique

DÉFINITION 14-3.36 *Soit $u \in \mathcal{S}(\mathbb{E}_n)$. L'application de $\mathbb{E}_n$ dans $\mathbb{R}$ définie par*

$$\Phi_u : \mathbb{E}_n \to \mathbb{R} \quad x \mapsto \Phi_u(x) = \langle u(x), x \rangle$$

est une forme quadratique, dite forme quadratique associée à l'endomorphisme symétrique u.

On vérifie facilement que $\varphi_u : (x, y) \mapsto \langle u(x), y \rangle$ est une forme bilinéaire symétrique (car $u \in \mathcal{S}(\mathbb{E}_n)$!) sur $\mathbb{E}_n$, et que Φ_u est la forme quadratique qui lui est associée.

THÉORÈME 14-3.37 L'application de $\mathcal{S}(\mathbb{E}_n)$ dans $\mathcal{Q}(\mathbb{E}_n)$ définie par

$$u \mapsto \Phi_u$$

est un isomorphisme d'espaces vectoriels : pour toute forme quadratique Φ sur l'espace $\mathbb{E}_n$, il existe un unique endomorphisme auto-adjoint de $\mathbb{E}_n$ tel que $\Phi = \Phi_u$. On dit que $u \in \mathcal{S}(\mathbb{E}_n)$ est associé à Φ.

Démonstration : Cette application est clairement linéaire. Comme les espaces $\mathcal{S}(\mathbb{E}_n)$ et $\mathcal{Q}(\mathbb{E}_n)$ ont même dimension $\frac{n(n+1)}{2}$, il suffit de prouver son injectivité. Si Φ_u est la forme nulle, il en est de même de sa forme polaire φ_u, et donc

$$\forall x, y \in \mathbb{E}_n \quad (u(x), y) = 0$$

ce qui montre bien que $u(x)$ est nul pour tout x de $\mathbb{E}_n$. ■

Lorsqu'on travaille dans une base orthonormale $\mathcal{B}$ de $\mathbb{E}_n$, la correspondance $u \mapsto \Phi_u$ se lit matriciellement

$$M(u, \mathcal{B}) = M(\Phi_u, \mathcal{B})$$

puisque, si $x, y \in \mathbb{E}_n$ sont représentés par les colonnes X et Y dans $\mathcal{B}$, et si M est la matrice de u dans cette base, on a

$$(u(x), y) = {}^t(MX)Y = {}^tX{}^tMY = {}^tXMY$$

puisque M est symétrique, et la matrice de φ_u dans $\mathcal{B}$ est donc bien M.

REMARQUE 14-3.38 Si on change de base orthonormale, $O \in \mathcal{O}_n(\mathbb{R})$ étant matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, on aura, d'après les formules de changement de bases

$$M' = M(u, \mathcal{B}') = O^{-1}MO$$
$$M'' = M(\Phi_u, \mathcal{B}) = {}^tOMO$$

et on aura encore $M' = M''$ car O est orthogonale.

REMARQUE 14-3.39 Dans une base quelconque où le produit scalaire est représenté par la matrice A qui appartient à $S_n^{*+}(\mathbb{R})$, on a

$$M(\Phi_u, \mathcal{B}) = {}^t[M(u, \mathcal{B})]A = AM(u, \mathcal{B})$$

EXEMPLE 14-3.40 Si $v \in \mathcal{L}(\mathbb{E}_n)$ est un endomorphisme quelconque, l'application $\Psi_v : \mathbb{E}_n \to \mathbb{R}$ définie par

$$\Psi_v(x) = \|v(x)\|^2$$

est clairement une forme quadratique positive sur $\mathbb{E}_n$. Comme

$$\Psi_v(x) = \langle v(x), v(x)\rangle = \langle v^*v(x), x\rangle \quad \text{avec} \quad v^*v \in \mathcal{S}(\mathbb{E}_n)$$

Ψ_v est associée à v^*v. Nous verrons ultérieurement que toute forme quadratique positive sur $\mathbb{E}_n$ est de ce type.

14-3.4.2 Théorème de réduction simultanée

Sur un espace euclidien, une forme quadratique est donc associée à un endomorphisme auto-adjoint. La diagonalisation de ce dernier va nous donner une réduction de la forme quadratique dans **une base orthonormale** de l'espace :

THÉORÈME 14-3.41 Soit Φ une forme quadratique sur un espace euclidien $(\mathbb{E}_n, \langle\ ,\ \rangle)$. Il existe une base orthonormale $\mathcal{B}$ de $\mathbb{E}_n$ dans laquelle la matrice de Φ est diagonale (base dite Φ-orthogonale). Toute base $\mathcal{B}$ vérifiant cette propriété est une base orthonormale de l'espace $\mathbb{E}_n$ diagonalisant l'endomorphisme symétrique associé à Φ.

Démonstration : Soient u l'endomorphisme symétrique associé à Φ, φ la forme polaire de Φ et $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base supposée répondre à la question. On a alors

$$\forall i \geqslant 2 \quad 0 = \varphi(e_1, e_i) = \langle u(e_1), e_i\rangle$$

ce qui montre que le vecteur $u(e_1)$ est orthogonal au sous-espace $\mathrm{vect}(e_i)_{i\geqslant 2}$. Ceci entraîne

$$u(e_1) \in \mathrm{vect}(e_1)$$

et e_1 est donc un vecteur propre de u. Comme il en est évidemment de même pour les autres vecteurs de $\mathcal{B}$, cette base est donc bien base orthonormale de vecteurs propres de u. Réciproquement, si $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ est base orthonormale qui diagonalise u, avec $(\lambda_i)_{1\leqslant i\leqslant n}$ la liste des valeurs propres (non nécessairement distinctes) associées, pour $x \in \mathbb{E}_n$ se décomposant sous la forme

$$x = \sum_{i=1}^{n} x_i e_i$$

l'expression de $\Phi(x)$

$$\Phi(x) = \langle u(x), x\rangle = \left\langle \sum_{i=1}^{n} \lambda_i x_i e_i, \sum_{i=1}^{n} x_i e_i \right\rangle = \sum_{i=1}^{n} \lambda_i x_i^2$$

montre que la base $\mathcal{B}$ est bien Φ-orthogonale. (on peut aussi dire que la matrice de Φ dans $\mathcal{B}$ est égale à celle de u, et vaut $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$).
■

Sur un espace vectoriel réel de dimension finie, se donner une structure euclidienne, c'est choisir une forme quadratique définie positive Ψ (que l'on décrète être le carré scalaire). Le théorème précédent peut donc aussi s'énoncer :

COROLLAIRE 14-3.42 Si Φ et Ψ sont deux formes quadratiques sur un espace vectoriel $\mathbb{E}$ réel de dimension finie, avec Ψ définie positive, il existe une base de $\mathbb{E}$ qui est à la fois orthonormale pour Ψ et orthogonale pour Φ.

On dit qu'une telle base réduit simultanément les formes quadratiques Φ et Ψ.

REMARQUE 14-3.43 Si Φ est une forme quadratique sur un espace euclidien $(\mathbb{E}_n, \langle\, , \rangle)$, et si l'expression de Φ **dans une base orthonormale** $\mathcal{B}_0$ de $\mathbf{E}_n$ est

$$\Phi(x) = \sum_{i=1}^{n} \alpha_i x_i^2$$

(où les x_i sont les coordonnées de x dans $\mathcal{B}_0$ et les α_i sont des réels indépendants de x), $\mathcal{B}_0$ est une base de réduction simultanée du carré scalaire et de Φ. La démonstration précédente montre que les $(\alpha_i)_{1\leqslant i\leqslant n}$ sont les valeurs propres de l'endomorphisme u symétrique associé à Φ. La signature de Φ est donc (p, q), avec p égal au nombre de valeurs propres >0 (comptées avec leurs multiplicités) et q étant le nombre de valeurs propres <0 de u. Si l'on connaît la signature de Φ, on connaît les signes des valeurs propres de u. Réciproquement, si le spectre de u est connu, la signature de Φ l'est aussi.

En particulier, si $A \in \mathcal{S}_n(\mathbb{R})$, on pourra obtenir des informations sur les signes des valeurs propres de A de la manière suivante : on travaille sur $\mathbf{R}^n$ muni de sa structure euclidienne canonique, et on considère la forme quadratique canoniquement associée à A

$$\Phi_A(X) = {}^tXAX$$

A est donc matrice dans la base canonique $\mathcal{B}_c$ (orthonormale) de l'endomorphisme symétrique associé à Φ_A. La signature de Φ_A (que l'on peut obtenir par exemple par la méthode de Gauss) détermine[19] les signes des valeurs propres de A.

14-3.5 Endomorphismes symétriques positifs, définis positifs

14-3.5.1 Définition

DÉFINITION 14-3.44 *Si u est un endomorphisme symétrique d'un espace euclidien $(\mathbb{E}_n, \langle\, , \rangle)$, on dit que u est positif si et seulement si la forme quadratique associée l'est :*

$$\forall x \in \mathbb{E}_n \quad \langle u(x), x\rangle \geqslant 0$$

On notera $\mathcal{S}^+(\mathbb{E}_n)$ l'ensemble de ces endomorphismes. De même, $\mathcal{S}^{*+}(\mathbb{E}_n)$ sera l'ensemble des endomorphismes symétriques définis positifs, c'est-à-dire tels que

$$\forall x \in \mathbf{E}_n \quad x \neq 0_{\mathbb{E}_n} \Rightarrow \langle u(x), x\rangle > 0$$

(un tel u est évidemment un automorphisme de $\mathbb{E}_n$).

[19] Ce raisonnement montre aussi que la signature d'une forme quadratique sur un espace réel $\mathbb{E}_n$ est liée aux signes des valeurs propres de sa matrice dans une base quelconque de $\mathbb{E}_n$. En particulier, deux matrices symétriques réelles congruentes ont des spectres distincts, mais les signes des valeurs propres de ces deux matrices se correspondent.

La matrice A de u dans une base orthonormale de $\mathbb{E}_n$ est symétrique, et on a évidemment

$$u \in \mathcal{S}^+(\mathbb{E}_n) \Leftrightarrow A \in \mathcal{S}_n^+(\mathbb{R})$$
$$u \in \mathcal{S}^{*+}(\mathbb{E}_n) \Leftrightarrow A \in \mathcal{S}_n^{*+}(\mathbb{R})$$

La remarque **14-3**.43 montre que, pour $u \in \mathcal{S}(\mathbb{E}_n)$

$$u \in \mathcal{S}^+(\mathbb{E}_n) \Leftrightarrow \text{ le spectre de } u \text{ est inclus dans } \mathbb{R}^+$$
$$u \in \mathcal{S}^{*+}(\mathbb{E}_n) \Leftrightarrow \text{ le spectre de } u \text{ est inclus dans } \mathbb{R}^{*+}$$

Nous avons vu (exemple **14-3**.40) que pour $v \in \mathcal{L}(\mathbb{E}_n)$, l'endomorphisme $u = v^*v$ est positif puisque

$$\forall x \in \mathbb{E}_n \quad \langle v^*v(x), x\rangle = \|v(x)\|^2$$

L'exercice qui suit montre qu'un endomorphisme symétrique positif peut toujours se factoriser sous cette forme :

EXERCICE 14-3.45 Soit u un endomorphisme symétrique positif d'un espace euclidien et $k \geqslant 2$ un entier. Montrer qu'il existe un unique $v \in \mathcal{S}^+(\mathbb{E}_n)$ tel que

$$v^k = u$$

(en particulier pour $k = 2$, u s'écrit $v^2 = v^*v$). On montrera d'abord que, si v existe, il laisse stable chaque sous-espace propre de u, et que l'endomorphisme induit par v sur chacun de ces sous-espaces est diagonalisable.

EXERCICE 14-3.46 Soient f et g deux endomorphismes d'un espace euclidien $\mathbb{E}_n$. Montrer que

$$f^*f = g^*g \Leftrightarrow \exists\, u \in \mathcal{O}(\mathbb{E}_n) \quad g = u \circ f$$

(on remarquera d'abord que, pour $x \in \mathbb{E}_n$, chacune de ces deux propriétés entraîne que $\|f(x)\| = \|g(x)\|$). Montrer que u est unique si et seulement si $f \in \mathcal{GL}(\mathbb{E}_n)$. En déduire que tout endomorphisme f de $\mathbb{E}_n$ peut se factorisersous la forme

$$f = u \circ s$$

avec u automorphisme orthogonal et s endomorphisme symétrique positif. A quelle condition cette écriture est-elle unique ?

EXERCICE 14-3.47 (DÉCOMPOSITION POLAIRE) L'exercice précédent a une traduction matricielle : toute matrice $A \in \mathcal{GL}_n(\mathbb{R})$ se factorise de manière unique sous la forme

$$A = OR$$

avec O orthogonale et R symétrique définie positive (décomposition "polaire" de la matrice A). On définit ainsi une bijection

$$\phi : \mathcal{O}_n(\mathbb{R}) \times \mathcal{S}_n^{*+}(\mathbb{R}) \to \mathcal{GL}_n(\mathbb{R})$$

Montrer que ϕ est bicontinue. Indication : pour la continuité de ϕ^{-1}, on utilisera la compacité de $\mathcal{O}_n(\mathbb{R})$.

14-3.5.2 Norme d'un endomorphisme d'un espace euclidien

Rappelons que, si u est un endomorphisme d'un espace vectoriel normé $(\mathbb{E}, \| \|)$ de dimension finie,

$$\|u\| = \sup_{x \neq 0_{\mathbb{E}}} \frac{\|u(x)\|}{\|x\|}$$

définit une norme sur $\mathcal{L}(\mathbb{E})$, dite norme subordonnée à la norme $\| \|$ sur $\mathbb{E}$. Lorsque l'espace est réel et que cette norme provient d'un produit scalaire $\langle .,. \rangle$, on a évidemment

$$\|u\|^2 = \sup_{x \neq 0_{\mathbb{E}}} \frac{\langle u(x), u(x) \rangle}{\langle x, x \rangle} = \sup_{\|x\|=1} \langle u(x), u(x) \rangle = \sup_{\|x\| \leqslant 1} \langle u(x), u(x) \rangle$$

On peut aussi exprimer cette norme par "dualité" :

PROPOSITION 14-3.48 Si u est un endomorphisme d'un espace euclidien $(\mathbb{E}_n, \langle\, ,\, \rangle)$, on a

$$\|u\| = \sup_{\substack{(x,y) \in \mathbb{E}_n \\ \|x\|=\|y\|=1}} \langle u(x), y \rangle$$

Démonstration : Par l'inégalité de Schwarz, on a pour $x, y \in \mathbb{E}_n$ de normes égales à 1

$$\langle u(x), y \rangle \leqslant \|u(x)\| \, \|y\| \leqslant \|u\| \, \|x\| \, \|y\| = \|u\|$$

ce qui montre que la borne supérieure considérée existe et vérifie

$$m = \sup_{\substack{(x,y) \in \mathbb{E}_n \\ \|x\|=\|y\|=1}} \langle u(x), y \rangle \leqslant \|u\|$$

Si x est un vecteur de norme 1 vérifiant $u(x) \neq 0_{\mathbb{E}_n}$, on a évidemment

$$\|u(x)\| = \left\langle u(x), \frac{u(x)}{\|u(x)\|} \right\rangle \leqslant m$$

et comme cette inégalité est encore vérifiée si $u(x)$ est nul, on obtient

$$\|u\| \leqslant m$$

ce qui démontre l'égalité souhaitée. ■

Comme pour $x, y \in \mathbb{E}_n$ on a $\langle u(x), y \rangle = \langle u^*(y), x \rangle$, on déduit de la formule précédente :

COROLLAIRE 14-3.49 Si u est un endomorphisme d'un espace euclidien

$$\|u\| = \|u^*\|$$

Lorsque l'endomorphisme u est symétrique, le théorème de réduction permet d'obtenir une autre expression de cette norme, et de la relier au spectre de u :

PROPOSITION 14-3.50 **Si** $u \in \mathcal{S}(\mathbb{E}_n)$, **on a**[20]

$$\|u\| = \sup_{\|x\| \leqslant 1} |\langle u(x), x\rangle| = \max\{|\lambda|, \lambda \in \mathrm{spec}(u)\}$$

Démonstration : Soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base orthonormale diagonalisant u, et $(\lambda_i)_{1 \leqslant i \leqslant n}$ les valeurs propres associées. Pour x de norme 1 décomposé dans cette base on a

$$\|u(x)\|^2 = \left\| u\left(\sum_{i=1}^n x_i e_i\right)\right\|^2 = \left\|\sum_{i=1}^n \lambda_i x_i e_i\right\|^2 = \sum_{i=1}^n \lambda_i^2 x_i^2$$

puisque la base est orthonormale. Si on pose

$$\rho = \max\{|\lambda|, \lambda \in \mathrm{spec}(u)\}$$

on a évidemment $\lambda_i^2 \leqslant \rho^2$ pour tout i, et donc

$$\|u(x)\|^2 \leqslant \rho^2 \left(\sum_{i=1}^n x_i^2\right) = \rho^2$$

avec égalité si on choisit pour x un des vecteurs e_j de la base pour lequel $|\lambda_j| = \rho$. On a donc bien

$$\|u\| = \rho = \max\{|\lambda|, \lambda \in \mathrm{spec}(u)\}$$

De plus, pour x de norme inférieure à 1, on a

$$|\langle u(x), x\rangle| = \left|\left\langle \sum_{i=1}^n \lambda_i x_i e_i, \sum_{i=1}^n x_i e_i\right\rangle\right| = \left|\sum_{i=1}^n \lambda_i x_i^2\right| \leqslant \rho \sum_{i=1}^n x_i^2 \leqslant \rho$$

avec égalité pour le même vecteur e_j que précédemment. Ceci donne bien

$$\rho = \sup_{\|x\| \leqslant 1} |\langle u(x), x\rangle| = \|u\| \quad \blacksquare$$

Si u est un endomorphisme quelconque de l'espace euclidien $\mathbb{E}_n$, en remarquant que

$$\forall x \in \mathbb{E}_n \quad \|u(x)\|^2 = \langle u^* u(x), x\rangle$$

et puisque u^*u est symétrique positif, on aura

COROLLAIRE 14-3.51 **Si** $u \in \mathcal{L}(\mathbb{E}_n)$, **on a**

$$\|u\| = \sqrt{\|u^* u\|} = \sqrt{\max(\mathrm{spec}(u^* u))}$$

En particulier, si on munit l'espace $\mathbb{R}^n$ de sa structure euclidienne canonique et de la norme associée, la norme matricielle subordonnée sur $\mathcal{M}_n(\mathbb{R})$ sera définie par

$$\forall A \in \mathcal{M}_n(\mathbb{R}) \quad \|A\| = \sqrt{\text{plus grande valeur propre de } {}^tAA}$$

[20] Lorque u est de plus positif, on obtiendra

$$\|u\| = \sup_{\|x\| \leqslant 1} \langle u(x), x\rangle = \max\{\lambda, \lambda \in \mathrm{spec}(u)\}$$

puisque toutes les quantités considérées sont positives.

14-3.6 Etude des automorphismes orthogonaux

14-3.6.1 Orientation d'un espace réel

Si $\mathbb{E}_n$ est un espace vectoriel réel de dimension finie >0, la relation binaire définie sur l'ensemble des bases de $\mathbb{E}_n$ par

$$\mathcal{B}_1 \equiv \mathcal{B}_2 \Leftrightarrow \det_{\mathcal{B}_1}(\mathcal{B}_2) > 0$$

est réflexive ($\det_{\mathcal{B}}(\mathcal{B}) = 1$), symétrique ($\det_{\mathcal{B}_2}(\mathcal{B}_1) = [\det_{\mathcal{B}_1}(\mathcal{B}_2)]^{-1}$) et transitive (utiliser la formule de changement de bases $\det_{\mathcal{B}_1}(\mathcal{B}_3) = \det_{\mathcal{B}_1}(\mathcal{B}_2)\det_{\mathcal{B}_2}(\mathcal{B}_3)$). Cette relation d'équivalence réalise une partition de l'ensemble des bases de $\mathbb{E}_n$ en deux classes d'équivalence : si $\mathcal{B}_0$ est une base arbitraire, il y a la classe de $\mathcal{B}_0$, et son complémentaire (non vide : on obtient une base qui n'est pas en relation avec $\mathcal{B}_0$ en changeant un des vecteurs de $\mathcal{B}_0$ en son opposé) qui est aussi une classe d'équivalence, puisque deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de ce complémentaire vérifient

$$\det_{\mathcal{B}_1}(\mathcal{B}_2) = \frac{\det_{\mathcal{B}_0}(\mathcal{B}_2)}{\det_{\mathcal{B}_0}(\mathcal{B}_1)} > 0$$

Orienter l'espace $\mathbb{E}_n$, c'est choisir une des ces deux classes et dire que toutes les bases composant cette classe sont directes (ou positives). Le complémentaire de cette classe est formé des bases indirectes (ou négatives). Il s'agit d'un choix arbitraire.

Si f est un automorphisme de $\mathbb{E}_n$, on a

$$\det_{\mathcal{B}}(f(\mathcal{B})) = \det f$$

On voit donc que les automorphismes de déterminant >0 sont ceux qui conservent l'orientation des bases (qui transforment les bases directes en bases directes). Permuter deux vecteurs d'une base, multiplier un vecteur par un réel négatif sont des opérations qui changent l'orientation d'une base. Faire opérer une permutation circulaire de longueur l sur les vecteurs d'une base ne change pas l'orientation si l est impair.

Dans le cadre d'un espace euclidien orienté, si $\mathcal{B}$ est une base orthonormale directe, une autre base orthonormale $\mathcal{B}'$ est directe si et seulement si la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice de $\mathcal{O}_n^+(\mathbb{R})$. Nous noterons $\mathcal{B}_{\mathrm{ON}}^+$ l'ensemble des bases orthonormales directes de l'espace.

14-3.6.2 Cas de la dimension 2

Nous travaillons ici dans un espace vectoriel $\mathbb{E}_2$ orienté, et nous fixons une base orthonormale directe $\mathcal{B}_0 = \left\{\vec{I}, \vec{J}\right\} \in \mathcal{B}_{\mathrm{ON}}^+$. Nous étudions les automorphismes orthogonaux de $\mathbb{E}_n$ en utilisant leur matrice dans la base $\mathcal{B}_0$:

- **Automorphismes orthogonaux directs :**

  Soit $u \in \mathcal{O}^+(\mathbb{E}_n)$. Sa matrice

$$M(u, \mathcal{B}_0) = A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

vérifie donc ${}^tAA = I_2$ et $\det A = 1$, soit

$$\begin{cases} a^2 + b^2 = c^2 + d^2 = 1 \\ ac + bd = 0 \\ ad - bc = 1 \end{cases}$$

Les deux premières égalités donnent l'existence de deux réels θ et ϕ (uniques modulo 2π) tels que

$$a = \cos\theta,\ b = \sin\theta,\ d = \cos\phi \text{ et } c = \sin\phi$$

La condition $ac + bd = 0$ donne $\sin(\theta + \phi) = 0$ et $ad - bc = 1$ s'écrit aussi $\cos(\theta + \phi) = 1$. On a donc $\phi = -\theta$ modulo 2π, et par conséquent

$$A = M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

où θ est un réel (unique modulo 2π). Réciproquement, toute matrice de ce type est dans $\mathcal{O}_2^+(\mathbb{R})$.

PROPOSITION 14-3.52 Le groupe multiplicatif $\mathcal{O}_2^+(\mathbb{R})$ est égal à

$$\mathcal{O}_2^+(\mathbb{R}) = \left\{ M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \theta \in \mathbb{R} \right\}$$

L'application

$$(\mathbb{R}, +) \to \left(\mathcal{O}_2^+(\mathbb{R}), \times\right) \quad \theta \mapsto M(\theta)$$

est un morphisme de groupes. $\left(\mathcal{O}_2^+(\mathbb{R}), \times\right)$ est commutatif.

Démonstration : Les formules d'addition pour les fonctions circulaires donnent immédiatement

$$M(\theta + \theta') = M(\theta)\, M(\theta')$$

pour tous θ et $\theta' \in \mathbb{R}$. ∎

PROPOSITION 14-3.53 Dans un espace euclidien orienté de dimension 2, la matrice d'un automorphisme orthogonal direct u dans une base orthonormale directe ne dépend pas de cette base. Il existe un réel θ (unique modulo 2π) tel que

$$\forall \mathcal{B} \in \mathcal{B}_{\mathrm{ON}}^+ \qquad M(u, \mathcal{B}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = M(\theta)$$

On dit que u est la rotation d'angle θ dans $\mathbb{E}_2$ orienté. Dans une base indirecte, la matrice de u est $M(-\theta)$: un changement de l'orientation de l'espace transforme l'angle d'une rotation en son opposé.

Démonstration : Dans la base $\mathcal{B}_0$, la matrice de u est de la forme $M(\theta)$. Si $\mathcal{B}$ est une autre base orthonormale directe de $\mathbb{E}_2$, la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}$ est dans $\mathcal{O}_2^+(\mathbb{R})$, donc de la forme $M(\varphi)$, avec $\varphi \in \mathbb{R}$. On a alors

$$M(u,\mathcal{B}) = M(\varphi)^{-1} M(\theta) M(\varphi) = M(\theta)$$

puisque les matrices commutent. La matrice de u est donc inchangée. Dans une base indirecte, θ sera changé en son opposé. Il suffit de faire la vérification dans une base indirecte particulière, par exemple $\mathcal{B}_0' = \left\{\vec{J}, \vec{I}\right\}$. On a

$$\begin{cases} u\left(\vec{J}\right) = \cos\theta\,\vec{J} - \sin\theta\,\vec{I} \\ u\left(\vec{I}\right) = \sin\theta\,\vec{J} + \cos\theta\,\vec{I} \end{cases}$$

ce qui donne bien

$$M(u,\mathcal{B}_0') = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = M(-\theta)$$

- **Automorphismes orthogonaux indirects.**

Si $v \in \mathcal{O}^-(\mathbb{E}_2)$, des calculs analogues à ceux qui précèdent montrent qu'il existe un réel θ tel que

$$M(v,\mathcal{B}_0) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

PROPOSITION 14-3.54 Les matrices orthogonales indirectes en dimension 2 forment l'ensemble

$$\mathcal{O}_2^-(\mathbb{R}) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \quad \theta \in \mathbb{R} \right\}$$

On a clairement $M(v,\mathcal{B}_0)^2 = I_2$, ce qui montre que tout élément v de $\mathcal{O}^-(\mathbb{E}_2)$ vérifie $v^2 = id_{\mathbb{E}_2}$, et est donc une involution. Comme v est de trace nulle, c'est une symétrie par rapport à une droite vectorielle, et comme v est un automorphisme orthogonal

$$\mathbb{E}_2 = \ker(v - id_{\mathbb{E}_2}) \overset{\perp}{\oplus} \ker(v + id_{\mathbb{E}_2})$$

et v est une symétrie orthogonale par rapport à une droite vectorielle (donc une réflexion puisqu'on est en dimension 2). On vérifie d'ailleurs aisément que, si

$$M(v,\mathcal{B}_0) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

le vecteur

$$\vec{u}\left(\frac{\theta}{2}\right) = \cos\frac{\theta}{2}\,\vec{I} + \sin\frac{\theta}{2}\,\vec{J}$$

est invariant par v, alors que

$$\vec{u}\left(\frac{\theta+\pi}{2}\right) = -\sin\frac{\theta}{2}\,\vec{I} + \cos\frac{\theta}{2}\,\vec{J}$$

est transformé en son opposé. L'égalité matricielle

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

traduit donc le fait qu'une rotation d'angle θ est produit de deux réflexions par rapport à deux droites vectorielles D_1 et D_2 telles que

$$(\widehat{D_1, D_2}) = \frac{\theta}{2} \mod \pi$$

PROPOSITION 14-3.55 Tout automorphisme orthogonal indirect en dimension 2 est une réflexion.

14-3.6.3 Cas de la dimension 3

En dimension 3, nous allons décrire un automorphisme orthogonal en montrant que, **dans une base orthonormale bien choisie**, sa matrice est "simple" :

- **Automorphismes orthogonaux directs :**

  On suppose $\mathbb{E}_3$ orienté. Si $u \in \mathcal{O}^+(\mathbb{E}_3)$, sa matrice M dans une base orthonormale quelconque est dans $\mathcal{O}_3^+(\mathbb{R})$. Elle possède au moins une valeur propre réelle (son polynôme caractéristique est de degré 3) égale à ± 1. En fait, 1 est nécessairement dans son spectre : si le spectre de u se réduisait à -1, et si $\vec{x}$ était vecteur propre associé, le plan $\mathbb{P} = (\vec{x})^\perp$ serait u-stable, l'endomorphisme v induit par u sur $\mathbb{P}$ vérifiant

  $$v \in \mathcal{O}(\mathbb{P}) \text{ et } \det v = -1$$

  Cet automorphisme orthogonal en dimension 2 serait donc une réflexion, et posséderait donc la valeur propre 1, qui serait dans le spectre[21] de u.

  Considérons donc un vecteur propre unitaire de u vérifiant

  $$u(\vec{e}_1) = \vec{e}_1$$

[21] Un calcul purement matriciel montre qu'en dimension impaire, une matrice orthogonale directe possède toujours 1 comme valeur propre : si $A \in \mathcal{O}_{2n+1}^+(\mathbb{R})$, on a

$$\det(A - I_{2n+1}) = \det({}^tA)\det(A - I_{2n+1}) = \det(I_{2n+1} - {}^tA)$$

ce qui donne facilement

$$\det(A - I_{2n+1}) = 0$$

Le plan vectoriel $\mathbb{P}_2$ orthogonal à $\vec{e}_1$ est alors orienté[22] par le choix d'une orientation sur sa normale, c'est-à-dire par le choix de ce vecteur $\vec{e}_1$. L'endomorphisme $v = u|_{\mathbb{P}_2}$ est évidemment dans $\mathcal{O}^+(\mathbb{P}_2)$, c'est donc une rotation d'angle θ dans $\mathbb{P}_2$ orienté : dans une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ orthonormale directe

$$M(u, \mathcal{B}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

u dite rotation d'axe $\operatorname{vect}(\vec{e}_1)$ orienté par $\vec{e}_1$ et d'angle θ (unique modulo 2π). Changer l'orientation de l'axe de la rotation change θ en son opposé.

PROPOSITION 14-3.56 Un automorphisme orthogonal direct u en dimension 3 est une rotation autour d'une droite vectorielle. Cette droite vectorielle est le sous-espace propre de u pour la valeur propre 1 (sauf lorsque $u = id_{\mathbb{E}_3}$).

Trouver l'axe d'une rotation donnée par sa matrice A dans une base orthonormale directe revient donc à la détermination du noyau de $A - I_3$. Une fois choisi un vecteur directeur X_1 qui oriente cet axe, il reste à déterminer l'angle θ. Comme la trace est un invariant de similitude, on a

$$\operatorname{trace} A = 2\cos\theta + 1$$

ce qui détermine $\pm\theta$ modulo 2π. Pour déterminer le signe correspondant à l'orientation de l'axe déterminée par X_1, on pourra faire un calcul de produit mixte : voir la section **14-4.1**.

REMARQUE 14-3.57 Lorsque $\theta = \pi$, u est alors la symétrie orthogonale par rapport à son axe $\operatorname{vect}(\vec{e}_1)$, qu'on appelle aussi retournement (ou demi-tour) d'axe $\operatorname{vect}(\vec{e}_1)$. L'égalité matricielle

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

montre que toute rotation est produit de deux retournements. Caractériser géométriquement ces demi-tours.

- **Automorphismes orthogonaux indirects :**

Si $u \in \mathcal{O}^-(\mathbb{E}_3)$, on montre cette fois que nécessairement $-1 \in \operatorname{spec}(u)$ (raisonnement analogue à celui qui précède). Si $\vec{e}_1$ est un vecteur propre associé, le plan orthogonal $\mathbb{P}_2 = (\vec{e}_1)^\perp$ est u-stable. On a $v = u|_{\mathbb{P}_2} \in \mathcal{O}^+(\mathbb{P}_2)$, et est

[22] On dit qu'une base orthonormale $(\vec{e}_2, \vec{e}_3)$ de $\mathbb{P}_2$ est directe si et seulement si $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est une base directe de $\mathbb{E}_3$. Prendre $-\vec{e}_1$ comme vecteur directeur de la normale à $\mathbb{P}_2$ change l'orientation de $\mathbb{P}_2$.

donc une rotation. Si $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est une base orthonormale directe, on trouvera donc un réel θ tel que

$$M(u, \mathcal{B}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

et on peut décrire u comme produit (commutatif) d'une rotation d'axe $\text{vect}(\vec{e}_1)$ et d'angle θ et d'une symétrie orthogonale par rapport au plan $(\vec{e}_1)^{\perp}$. On a cette fois (sauf lorsque $u = -id_{\mathbb{E}_3}$)

$$\text{vect}(\vec{e}_1) = \ker(u + id_{\mathbb{E}_3}) \text{ et } \operatorname{trace} u = 2\cos\theta - 1$$

14-3.6.4 Complément : réduction en dimension quelconque

L'étude menée en dimension 3 peut être généralisée :

EXERCICE 14-3.58 Si $u \in \mathcal{O}(\mathbb{E}_n)$, montrer qu'il existe une décomposition orthogonale de l'espace

$$\mathbb{E}_n = \mathbb{F}_1 \overset{\perp}{\oplus} \mathbb{F}_2 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} \mathbb{F}_m$$

en sous-espaces vectoriels de dimensions 1 ou 2 stables par u. Sur chaque sous-espace de dimension 1, u opère comme l'identité ou son opposée. Sur chaque sous-espace de dimension 2, u opère comme une rotation.

Indication : on peut d'abord se ramener au cas où le spectre de u est vide, en écrivant

$$\mathbb{E}_n = \ker(u - id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \ker(u + id_{\mathbb{E}_n}) \overset{\perp}{\oplus} \mathbb{S}$$

et en remarquant que $\mathbb{S}$ est alors u-stable, avec $u|_{\mathbb{S}} \in \mathcal{O}(\mathbb{S})$. Si le spectre de u est vide, considérer l'endomorphisme symétrique

$$v = u + u^* = u + u^{-1}$$

et montrer que si $x \in \mathbb{E}_n$ est un vecteur propre de v, le plan $\mathbb{F}_1 = \text{vect}(x, u(x))$ est u-stable. Montrer que $u|_{\mathbb{F}_1}$ est une rotation. Conclure enfin en remarquant que $\mathbb{F}_1^{\perp}$ est u-stable.

Toute matrice A de $\mathcal{O}_n(\mathbb{R})$ est donc semblable à une matrice diagonale par blocs de la forme

$$\begin{pmatrix} I_p & 0 & 0 & 0 & 0 \\ 0 & -I_q & 0 & 0 & 0 \\ 0 & 0 & \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \begin{pmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{pmatrix} \end{pmatrix}$$

(avec éventuellement p, q ou r nuls) et $\theta_i \in \mathbb{R} - \pi\mathbb{Z}$. Il est clair que $A \in \mathcal{O}_n^+(\mathbb{R})$ si et seulement si q est pair.

EXERCICE 14-3.59 En déduire que $\mathcal{O}_n^+(\mathbb{R})$ est connexe par arcs.

14-4 Rappels : produit mixte et produit vectoriel

14-4.1 Produit mixte

PROPOSITION 14-4.1 Dans un espace euclidien orienté $\mathbb{E}_n$, le déterminant d'une famille de n vecteurs $(x_1, \ldots, x_n)$ dans une base orthonormale directe $\mathcal{B}$ ne dépend pas de cette base. Ce déterminant est appelé produit mixte des vecteurs $(x_i)_{1 \leqslant i \leqslant n}$ et est noté

$$[x_1, \ldots, x_n] = \det_{\mathcal{B}} (x_1, \ldots, x_n)$$

Démonstration : Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthonormales directes, on a

$$\det_{\mathcal{B}} (\mathcal{B}') = 1$$

puisque la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est dans $\mathcal{O}_n^+ (\mathbb{R})$, et la formule de changement de base donne

$$\det_{\mathcal{B}'} (x_1, \ldots, x_n) = \det_{\mathcal{B}'} (\mathcal{B}) \det_{\mathcal{B}} (x_1, \ldots, x_n) = \det_{\mathcal{B}} (x_1, \ldots, x_n) \quad \blacksquare$$

Ce produit mixte possède donc toutes les propriétés du déterminant. Un changement d'orientation de l'espace le transforme en son opposé. Si u est un endomorphisme de $\mathbb{E}_n$, on a

$$[u (x_1), \ldots, u (x_n)] = \det u . [x_1, \ldots, x_n]$$

Il en résulte qu'un automorphisme de déterminant égal à 1 (en particulier un automorphisme orthogonal direct) conserve le produit mixte.

Si $(x_i)_{1 \leqslant i \leqslant n}$ est une famille libre, le produit mixte $[x_1, \ldots, x_n]$ est strictement positif si et seulement si la base $(x_i)_{1 \leqslant i \leqslant n}$ est directe. Quant à la valeur absolue de ce produit mixte (qui ne dépend pas de l'orientation de l'espace), nous l'interprèterons dans la section suivante comme le "volume" du parallélépipède construit sur les vecteurs $(x_i)_{1 \leqslant i \leqslant n}$. Ceci pourrait aussi se faire en traitant l'exercice

EXERCICE 14-4.2 Dans un espace euclidien orienté $\mathbb{E}_n$, montrer que

$$|[x_1, \ldots, x_n]| \leqslant \prod_{i=1}^{n} \|x_i\|$$

et traiter le cas de l'égalité. On pourra supposer les vecteurs indépendants et raisonner par récurrence sur n, en travaillant dans une base orthonormale dont le premier vecteur est $e_1 = \dfrac{x_1}{\|x_1\|}$. En déduire l'inégalité d'Hadamard

$$\forall A = (a_{i,j}) \in \mathcal{M}_n (\mathbb{R}) \qquad |\det A| \leqslant \prod_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}$$

REMARQUE 14-4.3 La notion de produit mixte permet de préciser l'angle d'une rotation : si $A \in \mathcal{O}_3^+(\mathbb{R})$ est la matrice d'une rotation dans une base orthonormale directe, un vecteur directeur de l'axe de la rotation est déterminé par la colonne X_1 de ses coordonnées dans cette base qui vérifie

$$AX_1 = X_1$$

On sait que l'angle θ de la rotation vérifie

$$\operatorname{trace} A = 2\cos\theta + 1$$

et $\pm\theta$ est déterminé modulo 2π. Si on choisit un vecteur X_2 orthogonal à X_1, on aura en complétant $\left(\frac{X_1}{\|X_1\|}, \frac{X_2}{\|X_2\|}\right)$ en une base orthonormale directe $\left(\frac{X_1}{\|X_1\|}, \frac{X_2}{\|X_2\|}, X_3\right)$

$$AX_2 = \cos\theta . X_2 + \|X_2\| \sin\theta . X_3$$

et le produit mixte $[X_1, X_2, AX_2]$ vaut

$$[X_1, X_2, AX_2] = \|X_1\| \, \|X_2\|^2 \sin\theta . [X_1, X_2, X_3] = \|X_1\| \, \|X_2\|^2 \sin\theta$$

et a donc le signe de $\sin\theta$. Si $[X_1, X_2, AX_2] > 0$, on choisira donc $\theta \in]0, \pi[$ déterminé par son cosinus. Si ce produit mixte est négatif, on choisira $\theta \in]-\pi, 0[$. Un peu de réflexion montre d'ailleurs qu'il n'est pas nécessaire de supposer X_2 orthogonal à X_1. Il suffit en fait de prendre X_2 non colinéaire à X_1.

14-4.2 Exercice : déterminants de Gram

Si $(u_i)_{1\leqslant i\leqslant p}$ sont p vecteurs d'un espace préhilbertien réel $\mathbb{E}$, on appelle matrice de Gram de ces p vecteurs la matrice de $\mathcal{S}_p(\mathbb{R})$ définie par

$$G(u_1, \ldots, u_p) = ((u_i, u_j))_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant p}}$$

Le déterminant de Gram des $(u_i)_{1\leqslant i\leqslant p}$ est

$$\Delta(u_1, \ldots, u_p) = \det G(u_1, \ldots, u_p) = \det((u_i, u_j))_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant p}}$$

Montrer que la forme quadratique sur $\mathbb{R}^p$ canoniquement associée à la matrice $G(u_1, \ldots, u_p)$ est

$$(x_1, \ldots, x_p) \mapsto \left\| \sum_{i=1}^{p} x_i u_i \right\|^2$$

En déduire que le déterminant de Gram $\Delta(u_1, \ldots, u_p)$ est toujours positif ou nul, et qu'il est strictement positif si et seulement si la famille $(u_i)_{1\leqslant i\leqslant p}$ est libre[23].

[23] On remarque donc que, dans un espace préhilbertien réel, la nullité d'un seul déterminant permet de conclure à la dépendance linéaire d'une famille de vecteurs.

Lorsque la famille $(u_i)_{1\leqslant i\leqslant p}$ est libre, on peut relier le déterminant de Gram à la notion de produit mixte :

Si $\mathcal{B}_p$ est une base orthonormale de $\text{vect}\,(u_i)_{1\leqslant i\leqslant p}$, que l'on utilise pour orienter l'espace engendré par ces vecteurs[24], en considérant le produit ${}^tMM$ où M est la matrice représentative de $(u_i)_{1\leqslant i\leqslant p}$ dans $\mathcal{B}_p$, montrer que

$$([u_1,\dots,u_p])^2 = \Delta\,(u_1,\dots,u_p)$$

En déduire qu'on ne change pas le déterminant de Gram si on permute les vecteurs $(u_i)_{1\leqslant i\leqslant p}$, ou si l'on rajoute à un des vecteurs une combinaison linéaire des autres. Si on décompose le vecteur u_1 en

$$u_1 = u_1' + u_1'' \quad \text{dans} \quad \mathbb{E} = \text{vect}\,(u_2,\dots,u_p) \overset{\perp}{\oplus} \text{vect}\,(u_2,\dots,u_p)^{\perp}$$

montrer que

$$\Delta\,(u_1,\dots,u_p) = \|u_1''\|^2\,\Delta\,(u_2,\dots,u_p)$$

et, en raisonnant par récurrence sur p, interpréter $\Delta\,(u_1,\dots,u_p)$ comme le carré du "volume" (en dimension p) du parallélépipède construit sur $(u_1,\dots,u_p)$. Montrer que la formule précédente permet aussi de déterminer la distance d'un vecteur à un sous-espace de dimension finie de $\mathbb{E}$, lorsque l'on connaît une base de ce sous-espace.

14-4.3 Produit vectoriel

Nous nous placerons ici dans le cas d'un espace euclidien $\mathbb{E}_3$ orienté de dimension 3 sur lequel, conformément à l'usage, nous noterons

$$\langle\vec{x},\vec{y}\rangle = \vec{x}.\vec{y}$$

le produit scalaire. Si $\vec{x}$ et $\vec{y}$ sont deux vecteurs de $\mathbb{E}_3$, l'application

$$\mathbb{E}_3 \to \mathbb{R} \quad \vec{z} \mapsto [\vec{x},\vec{y},\vec{z}]$$

est une forme linéaire sur $\mathbb{E}_3$. Il existe donc un unique vecteur qu'on appelle produit vectoriel des vecteurs $\vec{x}$ et $\vec{y}$, noté $\vec{x}\wedge\vec{y}$ (ou $\vec{x}\times\vec{y}$) tel que

$$\forall\,\vec{z}\in\mathbb{E}_3 \quad [\vec{x},\vec{y},\vec{z}] = (\vec{x}\wedge\vec{y})\,.\vec{z}$$

Le produit vectoriel est ainsi défini par dualité (il est clair qu'un changement d'orientation de l'espace transforme le produit vectoriel en son opposé). De nombreuses propriétés du produit vectoriel en découlent :

- L'application

  $$\mathbb{E}_3\times\mathbb{E}_3 \to \mathbb{E}_3 \quad (\vec{x},\vec{y}) \mapsto \vec{x}\wedge\vec{y}$$

  est bilinéaire antisymétrique.

  En effet, l'application $\mathbb{E}_3\times\mathbb{E}_3 \to \mathbb{E}_3^*$ définie par $(\vec{x},\vec{y}) \mapsto [\vec{x},\vec{y},\bullet]$ l'est clairement.

[24] Orientation qui n'a en l'occurence aucune importance, puisque les produits mixtes sont élevés au carré.

- Si $\vec{x}$ et $\vec{y} \in \mathbb{E}_3$, on a

$$\vec{x} \wedge \vec{y} = \vec{0}_{\mathbb{E}_3} \Leftrightarrow \{\vec{x}, \vec{y}\} \text{ liée}$$

Si $\{\vec{x}, \vec{y}\}$ est liée, on a clairement $[\vec{x}, \vec{y}, \bullet] = 0_{\mathbb{E}_3^*}$. Réciproquement, si $\{\vec{x}, \vec{y}\}$ est libre, on peut compléter en une base $\{\vec{x}, \vec{y}, \vec{z}\}$ de $\mathbb{E}_3$, et donc

$$[\vec{x}, \vec{y}, \bullet] \neq 0_{\mathbb{E}_3^*}$$

- Si $\{\vec{x}, \vec{y}\}$ est libre, $\vec{x} \wedge \vec{y} \neq \vec{0}_{\mathbb{E}_3}$ est un vecteur orthogonal au plan vect$(\vec{x}, \vec{y})$. La base $(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y})$ est directe, et $\|\vec{x} \wedge \vec{y}\|$ est égal à la surface du parallélogramme construit sur $\vec{x}$ et $\vec{y}$.

En effet, par construction

$$(\vec{x} \wedge \vec{y}) . \vec{x} = [\vec{x}, \vec{y}, \vec{x}] = 0 = [\vec{x}, \vec{y}, \vec{y}] = (\vec{x} \wedge \vec{y}) . \vec{y}$$

On a de plus $[\vec{x}, \vec{y}, \vec{x} \wedge \vec{y}] = \|\vec{x} \wedge \vec{y}\|^2 > 0$, ce qui montre que $(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y})$ est directe et que $\|\vec{x} \wedge \vec{y}\|^2$ est le volume du parallélépipède construit sur $(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y})$, dont la hauteur a pour longueur $\|\vec{x} \wedge \vec{y}\|$ et dont le parallélogramme de base est construit sur $\vec{x}$ et $\vec{y}$.

- Si $\vec{x} = x_1\vec{I} + x_2\vec{J} + x_3\vec{K}$ et $\vec{y} = y_1\vec{I} + y_2\vec{J} + y_3\vec{K}$ dans une base orthonormale directe $\left(\vec{I}, \vec{J}, \vec{K}\right)$, alors

$$\vec{x} \wedge \vec{y} = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} \vec{I} - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} \vec{J} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \vec{K}$$

Il suffit en effet de calculer $[\vec{x}, \vec{y}, \vec{z}]$ dans cette base, en développant suivant la troisième colonne et en identifiant le résultat au produit scalaire $(\vec{x} \wedge \vec{y}) . \vec{z}$. On a en particulier

$$\vec{I} \wedge \vec{J} = \vec{K}, \quad \vec{J} \wedge \vec{K} = \vec{I} \quad \text{et} \quad \vec{K} \wedge \vec{I} = \vec{J}$$

- Identité de Lagrange :

$$\|\vec{x} \wedge \vec{y}\|^2 + (\vec{x}.\vec{y})^2 = \|\vec{x}\|^2 \|\vec{y}\|^2$$

On peut faire cette vérification en calculant dans une base orthonormale directe "adaptée au problème" (comme nous le ferons pour la formule du double produit vectoriel). On peut aussi utiliser le déterminant de Gram :

$$\|\vec{x} \wedge \vec{y}\|^4 = [\vec{x}, \vec{y}, \vec{x} \wedge \vec{y}]^2 = \Delta(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y}) = \begin{vmatrix} \|\vec{x}\|^2 & \vec{x}.\vec{y} & 0 \\ \vec{x}.\vec{y} & \|\vec{y}\|^2 & 0 \\ 0 & 0 & \|\vec{x} \wedge \vec{y}\|^2 \end{vmatrix}$$

ce qui donne le résultat lorsque $\vec{x}$ et $\vec{y}$ sont indépendants, la vérification étant immédiate lorsqu'ils sont liés. Si $\vec{x}$ et $\vec{y}$ sont deux vecteurs non nuls

de $\mathbb{E}_3$, on appelle angle non orienté des vecteurs $\vec{x}$ et $\vec{y}$ l'unique réel $\theta \in [0, \pi]$ tel que

$$\cos\theta = \frac{\vec{x}.\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}$$

L'identité de Lagrange donne alors

$$\sin\theta = \frac{\|\vec{x} \wedge \vec{y}\|}{\|\vec{x}\|\,\|\vec{y}\|}$$

égalité qui donne à nouveau l'interprétation de $\|\vec{x} \wedge \vec{y}\|$ comme surface du parallélogramme construit sur $\vec{x}$ et $\vec{y}$.

- Formule du double produit vectoriel :

$$\vec{x} \wedge (\vec{y} \wedge \vec{z}) = (\vec{x}.\vec{z})\ \vec{y} - (\vec{x}.\vec{y})\ \vec{z}$$

Cette égalité est une évidence si $\vec{y}$ et $\vec{z}$ sont liés. Sinon le vecteur $\vec{x} \wedge (\vec{y} \wedge \vec{z})$, étant orthogonal à $\vec{y} \wedge \vec{z}$, est nécessairement dans le plan $\operatorname{vect}(\vec{y}, \vec{z})$. On trouve ses coordonnées en travaillant dans une base orthonormale adaptée : on choisit $\vec{I}$ et $\vec{J}$ formant une base orthonormale de $\operatorname{vect}(\vec{y}, \vec{z})$ tels que

$$\vec{y} = \|\vec{y}\|\,.\vec{I} \quad \text{et} \quad \vec{z} = \alpha\vec{I} + \beta\vec{J}$$

et $\vec{K} = \vec{I} \wedge \vec{J}$. On a alors, pour $\vec{x} = x_1\vec{I} + x_2\vec{J} + x_3\vec{K}$

$$\vec{y} \wedge \vec{z} = \beta\,\|\vec{y}\|\,\vec{K} \quad \text{et} \quad \vec{x} \wedge (\vec{y} \wedge \vec{z}) = \beta\,\|\vec{y}\|\,x_2\vec{I} - \beta\,\|\vec{y}\|\,x_1\vec{J}$$

ce qui donne le résultat, puisque

$$(\vec{x}.\vec{z})\ \vec{y} = (\alpha x_1 + \beta x_2)\,\|\vec{y}\|\ \vec{I} \text{ et } (\vec{x}.\vec{y})\ \vec{z} = x_1\,\|\vec{y}\|\left(\alpha\vec{I} + \beta\vec{J}\right)$$

EXERCICE 14-4.4 Identité de Jacobi : vérifier que

$$\vec{x} \wedge (\vec{y} \wedge \vec{z}) + \vec{y} \wedge (\vec{z} \wedge \vec{x}) + \vec{z} \wedge (\vec{x} \wedge \vec{y}) = \vec{0}_{\mathbb{E}_3}$$

EXERCICE 14-4.5 Division vectorielle : résoudre l'équation d'inconnue $\vec{x}$, avec $\vec{a}$ donné vérifiant $\vec{a} \neq \vec{0}_{\mathbb{E}_3}$

$$\vec{a} \wedge \vec{x} = \vec{b}$$

Indication : une condition nécessaire d'existence d'une solution est $\vec{a}.\vec{b} = 0$. L'équation est linéaire, et l'ensemble des solutions de l'équation homogène est $\operatorname{vect}(\vec{a})$. Trouver une solution particulière colinéaire à $\vec{a} \wedge \vec{b}$.

EXERCICE 14-4.6 Si u est un automorphisme orthogonal de $\mathbb{E}_3$, comparer

$$u(\vec{x} \wedge \vec{y}) \text{ et } u(\vec{x}) \wedge u(\vec{y})$$

EXERCICE 14-4.7 Endomorphisme antisymétrique d'un espace euclidien : une application d'un espace euclidien $\mathbb{E}_n$ dans lui-même est dite antisymétrique si et seulement si

$$\forall x, y \in \mathbb{E} \quad \langle u(x), y\rangle = -\langle x, u(y)\rangle$$

On sait (*cf.* section **14-3.1.3**) que cette propriété entraîne la linéarité de u et équivaut alors à $u = -u^*$. Si $\mathcal{A}(\mathbb{E}_n)$ est l'espace vectoriel des endomorphismes antisymétriques de $\mathbb{E}_n$, on a

$$\mathcal{L}(\mathbb{E}_n) = \mathcal{S}(\mathbb{E}_n) \oplus \mathcal{A}(\mathbb{E}_n)$$

Dans le cas particulier d'un espace euclidien orienté de dimension 3, montrer que u est un endomorphisme antisymétrique si et seulement s'il existe un vecteur $\vec{\Omega}$ de $\mathbb{E}_3$ tel que

$$u = \vec{\Omega} \wedge \bullet$$

et que cette écriture de u est alors unique (travailler avec la matrice de u dans une base orthonormale directe).

Dans le cas général, montrer qu'un endomorphisme antisymétrique n'a aucune valeur propre (réelle !) non nulle. Montrer que, pour $u \in \mathcal{A}(\mathbb{E}_n)$, le sous-espace $\mathbb{F} = (\ker u)^{\perp}$ est u-stable et que u induit sur ce sous-espace un endomorphisme antisymétrique v de spectre vide. Montrer que $w = v^2 \in \mathcal{S}(\mathbb{F})$, et que si x est un vecteur propre de w, la famille

$$\left(\frac{x}{\|x\|}, \frac{u(x)}{\|u(x)\|}\right)$$

est orthonormale et engendre un plan u-stable, dont l'orthogonal dans $\mathbb{F}$ est aussi stable. En déduire qu'il existe une base orthonormale de $\mathbb{E}_n$ où la matrice de u s'écrit

$$\begin{pmatrix} O_p & 0 & 0 & 0 \\ 0 & \begin{pmatrix} 0 & -\alpha_1 \\ \alpha_1 & 0 \end{pmatrix} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \begin{pmatrix} 0 & -\alpha_r \\ \alpha_r & 0 \end{pmatrix} \end{pmatrix}$$

Déduire enfin de cette étude qu'une matrice réelle antisymétrique n'est $\mathbb{R}$-diagonalisable que si elle est nulle, que son spectre est inclus dans $i.\mathbb{R}$, axe des imaginaires purs, et qu'elle est toujours $\mathbb{C}$-diagonalisable (nous reverrons ce résultat dans le chapitre suivant consacré aux espaces hermitiens).

EXERCICE 14-4.8 Détermination rapide de l'axe et de l'angle d'une rotation d'un espace de dimension 3 orienté.

Dans $\mathbb{E}_3$ orienté, montrer que la rotation d'axe $\mathrm{vect}(\vec{e}_1)$ (orienté par le vecteur **unitaire** $\vec{e}_1$) et d'angle θ est donnée par

$$r(\vec{x}) = (\vec{x}.\vec{e}_1)\,\vec{e}_1 + \cos\theta.[\vec{x} - (\vec{x}.\vec{e}_1)\,\vec{e}_1] + \sin\theta.[\vec{e}_1 \wedge \vec{x}]$$

et que la partie antisymétrique de r est donnée par

$$a(\vec{x}) = \frac{1}{2}(r - r^*)(\vec{x}) = \sin\theta.[\vec{e}_1 \wedge \vec{x}]$$

En déduire que, si $M \in \mathcal{O}_3^+(\mathbb{R})$, un examen attentif de sa trace et de la matrice $\frac{1}{2}(M - {}^tM)$ permet d'obtenir l'axe et l'angle de la rotation associée à M.

EXERCICE 14-4.9 Dérivée d'un repère orthonormal mobile.

$\mathbb{E}_3$ désigne toujours un espace euclidien orienté de dimension 3. Nous considérons ici un intervalle $\mathcal{I}$ de $\mathbb{R}$ et une application de classe $\mathcal{C}^1$ de $\mathcal{I}$ dans $(\mathbb{E}_3)^3$

$$\mathcal{I} \ni t \mapsto \left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$$

telle que, pour tout $t \in \mathcal{I}$, $\left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$ soit une base orthonormale de $\mathbb{E}_3$. Si $\mathcal{B}_0$ est une base orthonormale directe de référence, l'application (continue sur l'intervalle $\mathcal{I}$ à valeurs dans $\{\pm 1\}$) définie par $t \mapsto \det_{\mathcal{B}_0}\left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$ est constante. L'orientation de la base $\mathcal{B}_t = \left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$ ne change pas, et nous supposerons donc que

$$\forall t \in \mathcal{I} \quad \mathcal{B}_t \in \mathcal{B}^+_{\mathrm{ON}}(\mathbb{E}_3)$$

Soit $t_0 \in \mathcal{I}$ fixé. Considérons l'endomorphisme D_{t_0} de $\mathbb{E}_3$ défini par les images des vecteurs de $\mathcal{B}_{t_0}$

$$D_{t_0}\left(\vec{I}(t_0)\right) = \vec{I'}(t_0) = \left.\frac{d}{dt}\vec{I}(t)\right|_{t=t_0}$$
$$D_{t_0}\left(\vec{J}(t_0)\right) = \vec{J'}(t_0) = \left.\frac{d}{dt}\vec{J}(t)\right|_{t=t_0}$$
$$D_{t_0}\left(\vec{K}(t_0)\right) = \vec{K}'(t_0) = \left.\frac{d}{dt}\vec{K}(t)\right|_{t=t_0}$$

Cet endomorphisme peut être simplement défini de la manière suivante :

Si α,β et γ sont trois réels fixés,

$$\forall t \in \mathcal{I} \quad \vec{M}(t) = \alpha\vec{I}(t) + \beta\vec{J}(t) + \gamma\vec{K}(t)$$

désigne un vecteur "mobile" **dont les coordonnées dans $\mathcal{B}_t$ sont fixes**. On a

$$\forall t \in \mathcal{I} \quad \vec{M}'(t) = \alpha\vec{I'}(t) + \beta\vec{J'}(t) + \gamma\vec{K}'(t)$$

ce qui donne en t_0

$$D_{t_0}\left(\vec{M}(t_0)\right) = \vec{M}'(t_0)$$

Montrer que, pour tout $t_0 \in \mathcal{I}$, D_{t_0} est un endomorphisme antisymétrique de $\mathbb{E}_3$, et qu'il existe donc un vecteur $\vec{\Omega}(t_0) \in \mathbb{E}_3$ tel que

$$D_{t_0} = \vec{\Omega}(t_0) \wedge \bullet$$

Montrer que l'application $t \mapsto \vec{\Omega}(t)$ est continue sur $\mathcal{I}$. Le vecteur $\vec{\Omega}(t_0)$ est appelé vecteur de rotation instantanée du repère mobile à l'instant t_0. Nous venons de retrouver l'expression du champ des vitesses d'un solide en mouvement par rapport à un repère fixe : en effet, si nous considérons un tel solide, nous pouvons symboliser son mouvement en considérant le mouvement d'un point particulier $t \mapsto A(t)$ du solide et le mouvement d'un trièdre orthonormé d'origine $A(t)$ lié au solide, et donc une base orthonormale mobile $t \mapsto \left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$. Si une application $t \mapsto B(t)$ représente le mouvement d'un autre point lié au solide le vecteur

$$\vec{M}(t) = \overrightarrow{A(t)B(t)}$$

a des coordonnées fixes dans la base mobile $\left(\vec{I}(t), \vec{J}(t), \vec{K}(t)\right)$. On a alors, conformément à ce qui précède

$$\forall t \in \mathcal{I} \quad \frac{d}{dt}\vec{M}(t) = \vec{\Omega}(t) \wedge \vec{M}(t)$$

Si O est l'origine du repère fixe, on a $\vec{M}(t) = \overrightarrow{OB(t)} - \overrightarrow{OA(t)}$, ce qui donne par dérivation

$$\vec{V}_B(t) = \vec{V}_A(t) + \vec{\Omega}(t) \wedge \overrightarrow{AB}$$

14-5 Complément : utilisation de polynômes orthogonaux

Cette section, sous forme d'exercices, traite des méthodes de quadrature approchées dues à Gauss. On utilise la notion de polynômes orthogonaux. Dans toute cette section, $a \in \mathbb{R}\cup\{-\infty\}$ et $b \in \mathbb{R}\cup\{+\infty\}$ vérifient $a < b$ et $p :]a,b[\to \mathbb{R}^{*+}$ est une fonction continue strictement positive (dite fonction de poids) telle que, pour tout $n \in \mathbb{N}$, la fonction $t \mapsto t^n p(t)$ soit intégrable sur $]a,b[$. Nous détaillerons le cas particulier

$$]a,b[=]-1,1[\text{ et } p(t) = \frac{1}{\sqrt{1-t^2}}$$

pour lequel on peut expliciter les différents coefficients intervenant dans les formules, mais les résultats subsistent dans le cas général. On note $U_n[X]$ l'ensemble des polynômes normalisés de degré n

$$U_n[X] = \{P \in \mathbb{R}[X] \mid P(X) - X^n \in \mathbb{R}_{n-1}[X]\}$$

14-5.0.1 Polynômes orthogonaux

EXERCICE 14-5.1 Pour P et $Q \in \mathbb{R}[X]$, on pose

$$\langle P(X), Q(X)\rangle = \int_a^b P(t)Q(t)\,p(t)\,dt$$

1. Montrer qu'on définit ainsi un produit scalaire sur $\mathbb{R}[X]$, vérifiant

$$\forall P,Q,R \in \mathbb{R}[X] \quad \langle P(X)R(X), Q(X)\rangle = \langle P(X), Q(X)R(X)\rangle$$

2. Montrer qu'il existe une unique suite $(P_n)_{n\in\mathbb{N}}$ telle que

$$\forall n \quad P_n \in U_n[X] \quad \text{et} \quad \forall n \neq m \quad \langle P_n, P_m\rangle = 0$$

Les $(P_n)_{n\in\mathbb{N}}$ sont appelés polynômes orthogonaux sur $]a,b[$, pour le poids p.

3. Montrer $(P_n)_{n\leqslant p}$ forme une base de $\mathbb{R}_p[X]$.

4. Montrer que $P_{n+2} - XP_{n+1}$ est orthogonal à $\mathbb{R}_{n-1}[X]$, et en déduire l'existence de deux scalaires λ_n et μ_n tels que

$$P_{n+2} = (X + \lambda_n) P_{n+1} - \mu_n P_n$$

Montrer que

$$\mu_n = \frac{\langle P_{n+1}, P_{n+1}\rangle}{\langle P_n, P_n\rangle} > 0$$

5. Montrer que pour $n \geqslant 1$

$$\int_a^b P_n(t)\, dt = 0$$

En déduire que P_n possède au moins une racine d'ordre impair dans l'intervalle $]a, b[$. On suppose que P_n possède exactement m racines $(\alpha_k)_{1 \leqslant k \leqslant m}$ d'ordre impair sur $]a, b[$, avec $m < n$. Montrer que

$$\left\langle P, \prod_{k=1}^{m} (X - \alpha_k) \right\rangle \neq 0$$

et obtenir une contradiction. En déduire que P_n est scindé dans $\mathbb{R}[X]$, et possède n racines simples dans $]a, b[$.

14-5.0.2 Polynômes de Tchebicheff

EXERCICE 14-5.2 On étudie à présent le cas particulier où $p(t) = \left(1 - t^2\right)^{-\frac{1}{2}}$ sur $]-1, 1[$.

- Montrer que, pour tout $n \in \mathbb{N}$, il existe un unique polynôme T_n tel que

$$\forall \theta \in \mathbb{R} \quad T_n(\cos\theta) = \cos n\theta$$

On pourra trouver une relation de récurrence sur les (T_n) en simplifiant

$$\cos(n+2)\theta + \cos n\theta$$

- Si $(P_n)_{n \in \mathbb{N}}$ est la suite de polynômes orthogonaux pour p, montrer que

$$\forall n \geqslant 1 \quad P_n = 2^{1-n} T_n$$

et expliciter les scalaires λ_n et μ_n intervenant dans la relation de récurrence vue dans l'exercice précédent.

- Expliciter les racines $\alpha_n < \alpha_{n-1} < \cdots < \alpha_1$ du polynôme P_n et calculer $\langle P_n, P_n\rangle$.

La situation rencontrée dans cet exercice est très particulière : pour un poids p quelconque, on ne peut obtenir en général que des valeurs approchées des racines des polynômes P_n. On peut cependant montrer[25] que, dans le cas général, les racines de P_n séparent celles de P_{n+1}.

[25] En utilisant la relation de récurrence et le fait que $\mu_n > 0$.

14-5.0.3 Méthode de quadrature de Gauss

EXERCICE 14-5.3 On utilise les résultats des deux exercices qui précèdent pour obtenir une valeur approchée de l'intégrale

$$\int_{-1}^{1} \frac{f(t)}{\sqrt{1-t^2}} dt$$

et une majoration de l'erreur pour $f : [-1,1] \to \mathbb{R}$ suffisamment régulière.

- Montrer que, pour f continue sur $[-1,1]$, l'intégrale considérée existe.
- Montrer que, si $(\beta_i)_{1\leqslant i\leqslant n}$ sont n points deux à deux distincts de $[-1,1]$, il existe un unique n-uplet de scalaires $(\mu_i)_{1\leqslant i\leqslant n}$ tels que

$$\forall P \in \mathbb{R}_{n-1}[X] \quad \int_{-1}^{1} \frac{P(t)}{\sqrt{1-t^2}} dt = \sum_{i=1}^{n} \mu_i P(\beta_i) \quad (*)$$

(utiliser la théorie de la dualité).

- Soit P un polynôme de $\mathbb{R}_{2n-1}[X]$ tels que $P(\alpha_i) = 0$ pour $i = 1, \ldots, n$ (les α_i désignant comme précédemment les racines du polynôme T_n). Montrer que

$$\int_{-1}^{1} \frac{P(t)}{\sqrt{1-t^2}} dt = 0$$

En déduire l'existence d'un n-uplet de scalaires $(\lambda_i)_{1\leqslant i\leqslant n}$ tels que

$$\forall P \in \mathbb{R}_{2n-1}[X] \quad \int_{-1}^{1} \frac{P(t)}{\sqrt{1-t^2}} dt = \sum_{i=1}^{n} \lambda_i P(\alpha_i)$$

(le domaine de validité de la formule d'intégration (*) est donc fortement augmenté pour un bon choix des β_i). Si on note Π_k le polynôme

$$\Pi_k(X) = \frac{T_n(X)}{X - \alpha_k}$$

montrer que[26]

$$\forall k \in \{1, \ldots, n\} \quad \lambda_k = \frac{1}{\Pi_k^2(\alpha_k)} \int_{-1}^{1} \frac{\Pi_k^2(t)}{\sqrt{1-t^2}} dt > 0$$

[26]Le fait que tous les coefficients soient positifs rend la méthode peu sensible aux erreurs d'arrondi : si $\tilde{P}(\alpha_k)$ est une valeur approchée de $P(\alpha_k)$, on a

$$\left|\sum_{k=1}^{n} \lambda_k P(\alpha_k) - \sum_{k=1}^{n} \lambda_k \tilde{P}(\alpha_k)\right| \leqslant \left(\max_i \left|P(\alpha_i) - \tilde{P}(\alpha_i)\right|\right) \sum_{k=1}^{n} \lambda_k$$

avec ici

$$\sum_{k=1}^{n} \lambda_k = \int_{-1}^{1} \frac{dt}{\sqrt{1-t^2}} = \pi$$

- Etablir l'existence d'une constante $\mu \in \mathbb{R}$ telle que

$$\forall P \in \mathbb{R}_{2n}[X] \quad \int_{-1}^{1} \frac{P(t)}{\sqrt{1-t^2}} dt = \sum_{i=1}^{n} \lambda_i P(\alpha_i) + \mu \, P^{(2n)}(0)$$

Montrer que

$$\mu = \frac{\pi}{(2n)!\,2^{2n-1}}$$

- Les résultats qui précèdent ont une généralisation presque immédiate lorsque le poids p est changé. Dans le cas particulier étudié ici, on a une valeur explicite très simple des (λ_i) : montrer que, pour $1 \leqslant p \leqslant n-1$

$$\sum_{k=1}^{n} T_p(\alpha_k) = 1$$

et en déduire que

$$\forall k \in \{1, \dots, n\} \quad \lambda_k = \frac{\pi}{n}$$

Pour f continue sur $[-1,1]$, on utilisera donc la formule d'intégration approchée (exacte pour les polynômes de degré $< 2n$)

$$\int_{-1}^{1} \frac{f(t)}{\sqrt{1-t^2}} dt \approx \frac{\pi}{n} \sum_{k=1}^{n} f(\alpha_k)$$

EXERCICE 14-5.4 Estimation de l'erreur.

- On suppose que la fonction f est de classe $\mathcal{C}^{2n}$ sur $[-1,1]$. Montrer qu'il existe un unique polynôme $P \in \mathbb{R}_{2n-1}[X]$ vérifiant

$$\forall i \in \{1, \dots, n\} \quad f(\alpha_i) = P(\alpha_i) \quad \text{et} \quad f'(\alpha_i) = P'(\alpha_i)$$

- Montrer que

$$\forall x \in [-1,1] \quad |f(x) - P(x)| \leqslant \frac{T_n^2(x)}{(2n)!\,2^{2n-2}} \left\| f^{(2n)} \right\|_\infty$$

(voir la section **10-4**).

- En déduire la majoration

$$\left| \int_{-1}^{1} \frac{f(t)}{\sqrt{1-t^2}} dt - \frac{\pi}{n} \sum_{k=1}^{n} f\left(\cos \frac{2k-1}{2n}\pi\right) \right| \leqslant \frac{\pi}{(2n)!\,2^{2n-1}} \left\| f^{(2n)} \right\|_\infty$$

REMARQUE 14-5.5 Les méthodes de quadrature de Gauss pour calculer une approximation de

$$\int_a^b f(t)\, p(t)\, dt$$

nécessitent la connaissance des valeurs approchées des racines du polynôme P_n (les notations sont celles de l'exercice **14-5**.1). Dans la pratique, on obtiendra ces valeurs par des techniques d'analyse numérique matricielle :

EXERCICE 14-5.6 Montrer que les racines du polynôme P_n sont les opposées des valeurs propres de la matrice tridiagonale

$$\begin{pmatrix} \lambda_{n-2} & \sqrt{\mu_{n-2}} & 0 & 0 & 0 \\ \sqrt{\mu_{n-2}} & \lambda_{n-3} & \sqrt{\mu_{n-3}} & 0 & 0 \\ 0 & \sqrt{\mu_{n-3}} & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \lambda_0 & \sqrt{\mu_0} \\ 0 & 0 & 0 & \sqrt{\mu_0} & \alpha \end{pmatrix}$$

avec $P_1(X) = X + \alpha$. Retrouver ainsi le fait que P_n est scindé dans $\mathbb{R}[X]$. Diagonaliser en particulier la matrice

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{\sqrt{2}}{2} \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

14-6 Exercices

EXERCICE 14-6.1 Soient $a_0, a_1, \ldots, a_n$ $(n+1)$ réels. Dans $\mathbb{R}_n[X]$, montrer que

$$< P, Q > = \sum_{j=0}^{n} P^{(j)}(a_j) Q^{(j)}(a_j)$$

définit un produit scalaire. Montrer qu'il existe une base orthonormale $\{P_0, \ldots, P_n\}$ avec $\forall i \;\; d^\circ P_i = i$. Expliciter ces polynômes lorsque tous les a_i sont égaux.

EXERCICE 14-6.2 Soit $(f_i)_{1 \leqslant i \leqslant n}$ une famille de fonctions continues sur $[0,1]$ à valeurs réelles. On note $a_{i,j} = \displaystyle\int_0^1 f_i(t) f_j(t)\,dt$. Montrer que la famille est libre dans $C^0([0,1], \mathbb{R})$ si et seulement si la matrice $A = (a_{i,j})$ est inversible. Montrer qu'alors toutes les valeurs propres de A sont strictement positives. Montrer que la matrice $A = (\alpha_{i,j})$ où $\alpha_{i,j} = \dfrac{1}{i+j}$ a un déterminant strictement positif.

EXERCICE 14-6.3 Déterminer les éléments propres de la matrice $(a_i b_j + a_j b_i)$ de $M_n(\mathbb{R})$.

EXERCICE 14-6.4 A l'aide d'une transformation orthogonale, réduire la forme quadratique :

$$\Phi(X) = \sum_{i=1}^{n} x_i^2 + \sum_{i<j} x_i x_j$$

EXERCICE 14-6.5 Déterminer, dans la base canonique de $\mathbb{R}^4$ euclidien la matrice de la symétrie orthogonale par rapport au plan d'équations $x + y + z + t = 0$ et $x + 2y + 3z + 4t = 0$.

EXERCICE 14-6.6 Montrer que $M = \frac{1}{3}\begin{pmatrix} -2 & -1 & 2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$ est une matrice de rotation dont on déterminera l'angle et l'axe.

EXERCICE 14-6.7 Soit A symétrique dans $M_n(\mathbb{R})$.

1. Montrer que $A \in M_n(\mathbb{R})$ est définie positive ssi tous ses mineurs principaux sont strictement positifs (les mineurs principaux sont les déterminants obtenus en enlevant les p dernières lignes et colonnes de A pour p compris entre 0 et $n-1$). On pourra raisonner par récurrence sur n.

2. Plus précisément, si tous les mineurs principaux de A sont non nuls, et si on note $(\Delta_1, \Delta_2, \ldots, \Delta_n)$ la suite de ces mineurs, ($\Delta_n = \det A$), montrer que la signature de la forme quadratique associée à A est égale à $(n-p, p)$, où p est le nombre de changements de signe dans la suite $(1, \Delta_1, \Delta_2, \ldots, \Delta_n)$.

EXERCICE 14-6.8 Si A est symétrique définie positive et B symétrique, montrer que AB et BA sont diagonalisables.

EXERCICE 14-6.9 Soit $A \in O_n(\mathbb{R})$. Montrer que, si -1 n'est pas valeur propre de A, il existe Q antisymétrique telle que :

$$A = (I_n + Q)^{-1}(I_n - Q) = (I_n - Q)(I_n + Q)^{-1}$$

et qu'alors $A \in SO_n(\mathbb{R})$. Etudier la réciproque.

EXERCICE 14-6.10 Soit $A = (a_{i,j})$ une matrice orthogonale . Montrer que

$$\Big|\sum_{i,j} a_{i,j}\Big| \leqslant n$$

EXERCICE 14-6.11 Soit $f \in \mathcal{L}(E_n)$ avec E_n euclidien. Montrer qu'il existe une base orthonormale de E_n transformée par f en famille orthogonale.

EXERCICE 14-6.12 <u>**Endomorphismes normaux d'un espace euclidien**</u>
$(E, <, >)$ est un espace euclidien de dimension n. $f \in \mathcal{L}(E)$ est dit normal ssi $ff^* = f^*f$ (ex : f autoadjoint, antisymétrique, orthogonal).

1. Décrire géométriquement f si $n = 2$.

2. Montrer que f et f^* ont même spectre et que tout sous-espace propre de f est stable par f^*.

3. Montrer que E est somme directe orthogonale de sev stables par f et f^* de dimensions 1 ou 2. (Raisonner par récurrence sur la dimension de l'espace. Si le spectre de f est vide, considérer un facteur irréductible $Q(X) = X^2 - 2\alpha X + \beta$ du polynôme caractéristique de f. Montrer que $\ker Q(f) \neq \{0_E\}$ et est f^*-stable. Considérer dans ce sous-espace un vecteur propre de f^*f).

EXERCICE 14-6.13 Pour $\lambda \in \mathbb{R}$ et $X = (x_1, \cdots, x_n) \in \mathbb{R}^n$, on pose

$$N_\lambda(X) = \sqrt{\sum_{i=1}^{n} x_i^2 + 2\lambda \sum_{i<j} x_i x_j}$$

Donner une condition nécessaire et suffisante sur λ pour que N_λ soit une norme. Si $\lambda < \mu$ et N_λ et N_μ sont deux normes, déterminer

$$\sup_{X\neq 0} \frac{N_\lambda(X)}{N_\mu(X)} \text{ et } \inf_{X\neq 0} \frac{N_\lambda(X)}{N_\mu(X)}$$

EXERCICE 14-6.14 $M_n(\mathbb{R})$ est muni de la norme matricielle associée à la norme euclidienne canonique sur $\mathbb{R}^n$. On note Φ l'ensemble des matrices $A \in M_n(\mathbb{R})$ vérifiant

$$\|A\| \leqslant 1 \text{ et } \forall B, C \quad (\|B\| \leqslant 1 \text{ et } \|C\| \leqslant 1 \text{ et } A = \frac{B+C}{2} \implies B = C = A)$$

1. Montrer que $A \in \Phi \implies \|A\| = 1$.
2. Montrer que $O_n(\mathbb{R}) \subset \Phi$.
3. En utilisant la décomposition polaire, montrer que l'inclusion précédente est en fait une égalité.

EXERCICE 14-6.15 On considère l'espace E des fonctions f continues de $\mathbb{R}$ dans $\mathbb{R}$ telles que l'intégrale $\int_{-\infty}^{+\infty} f^2(t)e^{-t^2}dt$ existe.

1. Montrer que E est un sous-espace vectoriel de $C^0(\mathbb{R}, \mathbb{R})$.
2. Montrer que $< f, g >= \int_{-\infty}^{+\infty} f(t)g(t)e^{-t^2}dt$ définit sur E un produit scalaire.
3. Montrer que la suite $(h_n)_{n\in\mathbb{N}}$ définie par $h_n(x) = e^{x^2}\frac{d^n}{dx^n}(e^{-x^2})$ est une famille orthogonale pour ce produit scalaire. Déterminer une suite réelle (c_n) pour que la famille $(c_n h_n)$ soit orthonormale.

EXERCICE 14-6.16 Soient A et B deux matrices symétriques réelles, q_A et q_B les formes quadratiques sur $\mathbb{R}^n$ qui leur sont canoniquement associées. On suppose que $\forall x \in \mathbb{R}^n \quad q_A(x) \geqslant q_B(x) \geqslant 0$. Montrer que $\det A \geqslant \det B$. Etudier les cas d'égalité.

EXERCICE 14-6.17 Soient a et b deux vecteurs d'un espace euclidien orienté de dimension 3. Résoudre le système

$$\begin{cases} < a, x > &= < b, y > \\ < a, y > &= < b, x > \\ a \wedge x &= b \wedge y \\ a \wedge y &= b \wedge x \end{cases}$$

EXERCICE 14-6.18 Démontrer que, si l'on ajoute à une forme quadratique positive le carré d'une forme linéaire, son discriminant dans une base fixée augmente.

EXERCICE 14-6.19 Soient A et B deux matrices symétriques réelles de taille n, dont le spectre est inclus dans $\mathbb{R}^+$.

1. Montrer que $\det(A + B) \geqslant \det A + \det B$.
2. Montrer que $(\det A)^{1/n} \leqslant (\text{tr}A)/n$.
3. Montrer que $(\det A)^{1/n} + (\det B)^{1/n} \leqslant (\det(A + B))^{1/n}$

EXERCICE 14-6.20 Soit $A \in \mathcal{M}_n(\mathbb{R})$. $\| \ \|$ est la norme euclidienne standard sur $\mathbb{R}^n$. Si $A \neq 0$ et $X \in \mathbb{R}^n$ avec $\|X\| = 1$, montrer que

$$\|AX\|^2 \geqslant (\det A)^2 \left(\frac{n-1}{\mathrm{trace}({}^tAA)}\right)^{n-1}$$

EXERCICE 14-6.21 Soit $A \in \mathcal{M}_{p,q}(\mathbb{R})$. Montrer que A et ${}^tAA$ ont même rang.

EXERCICE 14-6.22 Soit E espace euclidien, et soient f et g deux endomorphismes symétriques de E. On suppose qu'il existe une base B dans laquelle les matrices de f et g sont triangulaires supérieures. Montrer que $f \circ g = g \circ f$.

Chapitre 15

Espaces préhilbertiens complexes et hermitiens

15-1 Espace préhilbertien complexe

15-1.1 Application semilinéaire

DÉFINITION 15-1.1 *Si* $\mathbb{E}$ *et* $\mathbb{F}$ *sont deux espaces vectoriels complexes, une application* $\varphi : \mathbb{E} \to \mathbb{F}$ *est semi-linéaire si et seulement si*

$$\forall x, y \in \mathbb{E} \quad \varphi(x+y) = \varphi(x) + \varphi(y)$$

$$\forall x \in \mathbb{E} \quad \forall \lambda \in \mathbb{C} \quad \varphi(\lambda x) = \overline{\lambda}\varphi(x)$$

Beaucoup de propriétés des applications linéaires se retrouvent pour les applications semi-linéaires : images directes et réciproques de s.e.v, noyau, rang, etc.... On peut aussi représenter matriciellement une application semi-linéaire lorsque $\mathbb{E}$ et $\mathbb{F}$ sont de dimensions finies, et rapportés à des bases respectives $\mathcal{B}$ et $\mathcal{B}'$. La définition de la matrice est la même que pour les applications linéaires. Avec $A = M(\varphi, \mathcal{B}, \mathcal{B}')$, il faut noter que, si $x \in \mathbb{E}$ est représenté par la colonne X de ses coordonnées dans $\mathcal{B}$, son image $y = \varphi(x)$ a pour coordonnées dans $\mathcal{B}'$

$$Y = A\overline{X}$$

Attention ! La composée de deux applications semi-linéaires est linéaire.

15-1.2 Produit scalaire

15-1.2.1 Définition

DÉFINITION 15-1.2 *Soit* $\mathbb{E}$ *un* $\mathbb{C}$*-espace vectoriel. Une application* $\varphi : \mathbb{E}\times\mathbb{E} \to \mathbb{C}$ *est un produit scalaire sur* $\mathbb{E}$ *si et seulement si*

$$\forall x \in \mathbb{E} \quad \varphi(x,\bullet) \text{ est une forme linéaire sur } \mathbb{E}$$
$$\forall x,y \in \mathbb{E} \quad \varphi(y,x) = \overline{\varphi(x,y)}$$
$$\forall x \in \mathbb{E} \quad x \neq 0_{\mathbb{E}} \Rightarrow \Phi(x) = \varphi(x,x) \in \mathbb{R}^{*+}$$

On note souvent, comme dans le cas réel, (x,y) le produit scalaire des vecteurs x et y de $\mathbb{E}$, en changeant évidemment de notation lorsqu'on considère simultanément plusieurs produits scalaires sur un même espace.

REMARQUE 15-1.3 Les deux premières propriétés entraînent évidemment que

$$\forall x \in \mathbb{E} \quad \varphi(\bullet,x) \text{ est une forme semi-linéaire sur } \mathbb{E}$$

φ est donc **linéaire à droite et semi-linéaire à gauche.** On exprime parfois ceci en disant que φ est une forme sesquilinéaire sur $\mathbb{E}$. On traduit la seconde propriété en disant que φ **possède la symétrie hermitienne**. L'application $\Phi : x \mapsto \varphi(x,x)$ est aussi appelée **forme quadratique hermitienne associée** à φ. Comme

$$\varphi(x,x) = \overline{\varphi(x,x)}$$

la symétrie hermitienne entraîne qu'un vecteur x quelconque de $\mathbb{E}$ vérifie

$$\Phi(x) \in \mathbb{R}$$

La troisième propriété

$$\forall x \in \mathbb{E} - \{0_{\mathbb{E}}\} \quad \Phi(x) > 0$$

s'exprime par : **la forme quadratique hermitienne associée à φ est définie positive.**

EXEMPLE 15-1.4 On peut dans le cas complexe, redonner des exemples analogues à ceux étudiés dans le cas des espaces préhilbertiens réels :

- Le produit scalaire canonique sur $\mathbb{C}^n$ est donné par

$$(x,y) = \sum_{i=1}^{n} \overline{x_i} y_i$$

- Sur l'espace $l^2(\mathbb{N})$ des suites complexes de carré sommable, pour des suites $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$, la formule

$$(u,v) = \sum_{n=0}^{+\infty} \overline{u_n} v_n$$

  définit un produit scalaire.

- L'analogue "continu" de l'exemple précédent est le produit scalaire défini sur $\mathcal{C}^0([0,1],\mathbb{C})$ par

$$\langle f,g\rangle = \int_0^1 \overline{f(t)}g(t)\,dt$$

Lorsqu'on calcule avec un produit scalaire, il faut tenir compte de la semi-linéarité à gauche. Par exemple, on vérifie aisément que, pour x et $y \in \mathbb{E}$ et $\alpha, \beta \in \mathbb{C}$

$$\Phi(\alpha x + \beta y) = |\alpha|^2 \Phi(x) + 2\,\mathrm{Re}(\overline{\alpha}\beta\varphi(x,y)) + |\beta|^2 \Phi(y)$$

On en déduit facilement la formule de "polarisation"

$$\forall x, y \in \mathbb{E} \quad \varphi(x,y) = \frac{1}{4}[\Phi(x+y) - \Phi(x-y) - i\,\Phi(x+iy) + i\,\Phi(x-iy)]$$

15-1.2.2 Inégalité de Schwarz, norme hermitienne

PROPOSITION 15-1.5 Si φ est un produit scalaire sur le $\mathbb{C}$-ev $\mathbb{E}$, et si Φ est la forme quadratique hermitienne associée, on a

$$\forall x, y \in \mathbb{E} \quad |\varphi(x,y)|^2 \leqslant \Phi(x)\Phi(y)$$

avec égalité si et seulement si x et y sont liés.

Démonstration : On remarquera qu'il est ici indispensable de mettre le module dans le terme de gauche, sous peine d'écrire des inégalités entre nombres complexes. La démonstration est pratiquement la même que dans le cas réel : l'inégalité est évidente si x est nul ou si $\varphi(x,y) = 0$. Sinon, on calcule, pour $t \in \mathbb{C}$

$$\Phi(tx+y) = |t|^2 \Phi(x) + 2\,\mathrm{Re}\left(\bar{t}\varphi(x,y)\right) + \Phi(y) \geqslant 0$$

et, en écrivant sous forme trigonométrique

$$\varphi(x,y) = \rho e^{i\theta}$$

avec $\theta \in \mathbb{R}$ et $\rho > 0$, on évalue l'expression précédente pour $t = \lambda e^{i\theta}$ avec $\lambda \in \mathbb{R}$. On a alors

$$\forall \lambda \in \mathbb{R} \quad \lambda^2\Phi(x) + 2\lambda\rho + \Phi(y) \geqslant 0$$

et on conclut comme dans le cas réel. ■

COROLLAIRE 15-1.6 Si $\langle\,,\rangle$ est un produit scalaire sur $\mathbb{E}$,

$$\|x\| = \sqrt{\langle x,x\rangle}$$

définit une norme sur $\mathbb{E}$ dite norme (hermitienne) associée au produit scalaire.

Démonstration : Seule l'inégalité triangulaire (dite inégalité de Minkowski) nécessite une justification. Elle se démontre, comme sur $\mathbb{R}$, à l'aide de l'inégalité de Schwarz. On vérifie aisément que, pour x et $y \in \mathbb{E}$

$$\|x+y\| = \|x\| + \|y\| \Leftrightarrow \mathrm{Re}\left(\langle x,y\rangle\right) = \|x\|\,\|y\|$$

ce qui montre d'abord qu'il y a égalité dans l'inégalité de Schwarz, puis que x et y sont $\mathbb{R}^+$-proportionnels. ■

L'inégalité de Schwarz s'écrit donc

$$\forall x, y \in \mathbb{E} \quad |\langle x,y\rangle| \leqslant \|x\|\,\|y\|$$

et on vérifie que cette propriété entraîne **la continuité du produit scalaire** de $\mathbb{E} \times \mathbb{E}$ dans $\mathbb{C}$, lorsque $\mathbb{E}$ est muni de sa norme hermitienne. De même, l'identité de polarisation s'écrit

$$\forall x, y \in \mathbb{E} \quad \langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 - i\,\|x+iy\|^2 + i\,\|x-iy\|^2\right)$$

On obtient également **l'identité du parallélogramme** :

$$\forall x, y \in \mathbb{E} \quad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

On remarquera aussi que, pour un vecteur $x \neq 0_{\mathbb{E}}$, il existe une infinité de vecteurs unitaires colinéaires à x : ce sont tous les vecteurs de la forme

$$y = e^{i\theta}\frac{x}{\|x\|} \quad \text{avec } \theta \in \mathbb{R}$$

15-1.2.3 Orthogonalité

Deux vecteurs x et y sont orthogonaux ($x \perp y$) si $\langle x,y\rangle = 0$. Cette relation est évidemment symétrique. On a toujours la relation de Pythagore

$$x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

Il faut noter que la réciproque est fausse dans le cas complexe :

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow \langle x,y\rangle \in i\,\mathbb{R}$$

En particulier, cette relation est toujours vérifiée pour $y = i\,x$.

Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}$, on note

$$\mathbb{F}^{\perp} = \{x \in \mathbb{E} \mid \forall y \in \mathbb{F} \quad \langle x,y\rangle = 0\}$$

c'est un sous-espace vectoriel de $\mathbb{E}$. Comme dans le cas réel, on a, pour $\mathbb{F}$ et $\mathbb{G}$ sous-espaces de $\mathbb{E}$

$$\begin{cases} \mathbb{F} \cap \mathbb{F}^{\perp} = \{0_{\mathbb{E}}\} \\ (\mathbb{F}+\mathbb{G})^{\perp} = \mathbb{F}^{\perp} \cap \mathbb{G}^{\perp} \\ \mathbb{F} \subset \left(\mathbb{F}^{\perp}\right)^{\perp} \\ \mathbb{F}^{\perp} + \mathbb{G}^{\perp} \subset (\mathbb{F} \cap \mathbb{G})^{\perp} \end{cases}$$

PROPOSITION 15-1.7 Si x est un vecteur non nul

$$x^{\perp} = \ker\langle x, \bullet\rangle$$

est un hyperplan fermé de $(\mathbb{E}, \|\ \|)$

15-1.3 Projection orthogonale

PROPOSITION 15-1.8 Soient $\mathbb{F}$ un sous-espace de $\mathbb{E}$, et $a \in \mathbb{E}$. Il existe $a_1 \in \mathbb{F}$ tel que

$$\|a - a_1\| = d(a, \mathbb{F}) = \inf\{\|a - x\|, x \in \mathbb{F}\}$$

si et seulement si $a \in \mathbb{F} \oplus \mathbb{F}^{\perp}$. Le vecteur a_1 est alors unique, il est égal à la composante de a selon $\mathbb{F}$ dans la somme directe $\mathbb{F} \oplus \mathbb{F}^{\perp}$.

Démonstration : Comme dans le cas réel, on montre que la condition est suffisante en utilisant la relation de Pythagore. Réciproquement, si a_1 existe, on a

$$\begin{aligned} \forall x \in \mathbb{F} \quad \forall \lambda \in \mathbb{C} \quad & \|a - (a_1 + \lambda x)\|^2 \\ &= \|a - a_1\|^2 + |\lambda|^2 \|x\|^2 - 2\,\mathrm{Re}\,(\lambda \langle a - a_1, x \rangle) \geqslant \|a - a_1\|^2 \end{aligned}$$

ce qui permet facilement de prouver que $\langle a - a_1, x \rangle = 0$ pour $x \in \mathbb{F}$ quelconque. ■

Lorsque $\mathbb{F}$ est de dimension finie, l'existence de bases orthonormales va nous permettre de montrer que tout $a \in \mathbb{E}$ possède une projection orthogonale sur $\mathbb{F}$:

15-1.4 Bases orthonormales en dimension finie

Comme dans le cas réel, une famille $(u_i)_{i \in I}$ est orthonormale si et seulement si

$$\forall i \neq j \quad \langle u_i, u_j \rangle = 0 \quad \text{et} \quad \forall i \in I \quad \|u_i\| = 1$$

Une telle famille est libre.

15-1.4.1 Existence. Projection sur un s.e.v de dimension finie

PROPOSITION 15-1.9 Si $\mathbb{E}_n$ est un espace préhilbertien complexe de dimension finie $n > 0$, il possède des bases orthonormales.

Démonstration : Par récurrence sur n. Si $n = 1$, un vecteur $x \neq 0_{\mathbb{E}}$ donne un vecteur

$$e_1 = \frac{x}{\|x\|}$$

qui forme une base "orthonormale" de $\mathbb{E}_1$. Si on suppose le résultat établi en dimension $n - 1$, on considère x et $e_1 \in \mathbb{E}_n$ comme précédemment. L'hyperplan $e_1^{\perp}$ vérifie alors

$$e_1 \notin e_1^{\perp} \Rightarrow \mathbb{E}_n = \mathrm{vect}\,(e_1) \overset{\perp}{\oplus} e_1^{\perp}$$

et on applique l'hypothèse de récurrence à $e_1^{\perp}$ muni de la restriction du produit scalaire, que l'on considère comme espace préhilbertien de dimension $n - 1$. Si $(e_k)_{2 \leqslant k \leqslant n}$ en est une base orthonormale, $(e_i)_{1 \leqslant i \leqslant n}$ est base orthonormale de $\mathbb{E}_n$. ■

Dans une telle base, les expressions du produit scalaire et de la norme hermitienne sont simples :

$$x = \sum_{i=1}^{n} x_i e_i \quad \text{et} \quad y = \sum_{i=1}^{n} y_i e_i \Rightarrow \langle x, y \rangle = \sum_{i=1}^{n} \overline{x_i} y_i \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^{n} |x_i|^2$$

Comme on a aussi $x_i = \langle e_i, x \rangle$ (on notera que x est écrit à droite dans ce produit scalaire), ceci peut s'écrire également

$$\langle x, y \rangle = \sum_{i=1}^{n} \overline{\langle e_i, x \rangle} \langle e_i, y \rangle \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^{n} |\langle e_i, x \rangle|^2$$

COROLLAIRE 15-1.10 Si $\mathbb{F}_p$ est un sous-espace de dimension finie d'un espace préhilbertien complexe, on a

$$\mathbb{E} = \mathbb{F}_p \oplus \mathbb{F}_p^\perp$$

Si $(e_i)_{1 \leqslant i \leqslant p}$ est une base orthonormale de $\mathbb{F}_p$, la projection orthogonale sur $\mathbb{F}_p$ d'un vecteur $x \in \mathbb{E}$ quelconque est donnée par

$$p_{\mathbb{F}_p}(x) = \sum_{i=1}^{p} \langle e_i, x \rangle e_i$$

Démonstration : On remarquera que cette projection, dépendant linéairement du vecteur x, fait intervenir ce vecteur à droite dans les produits scalaires. On vérifie simplement que

$$\forall j \in \{1, \ldots, p\} \quad \left\langle e_j, x - \sum_{i=1}^{p} \langle e_i, x \rangle e_i \right\rangle = 0$$

ce qui montre bien que

$$x = p_{\mathbb{F}_p}(x) + [x - p_{\mathbb{F}_p}(x)] \in \mathbb{F}_p \oplus \mathbb{F}_p^\perp \quad \blacksquare$$

15-1.4.2 Procédé de Schmidt

Comme dans le cas réel, on a

THÉORÈME 15-1.11 Soit $(u_i)_{1 \leqslant i \leqslant p}$ une famille libre d'un espace préhilbertien complexe $\mathbb{E}$. Il existe une famille orthonormale unique $(e_i)_{1 \leqslant i \leqslant p}$ de l'espace $\mathbb{E}$ telle que

$$\forall i \in \{1, \ldots, p\} \quad \text{vect}(e_1, \ldots, e_i) = \text{vect}(u_1, \ldots, u_i)$$
$$\forall i \in \{1, \ldots, p\} \quad \langle e_i, u_i \rangle \in \mathbb{R}^{*+}$$

Cette famille peut être construite de proche en proche par l'algorithme de Gram-Schmidt :

$$e_1 = \frac{u_1}{\|u_1\|} \text{ et pour } i \geqslant 2 \quad e_i = \frac{u'_i}{\|u'i\|} \text{ avec } u'_i = u_i - \sum_{k=1}^{i-1} \langle e_k, u_i \rangle e_k$$

La démonstration est en tout point analogue à celle du chapitre précédent.

15-1.4.3 Inégalité de Bessel

PROPOSITION 15-1.12 Si $(e_n)_{n\in\mathbb{N}}$ est une suite orthonormale d'un espace préhilbertien, pour tout $x \in \mathbb{E}$, la suite complexe $(\langle e_n, x\rangle)_{n\in\mathbb{N}}$ est de carré sommable et

$$\sum_{n=0}^{+\infty} |\langle e_n, x\rangle|^2 \leqslant \|x\|^2$$

Il y a égalité dans cette inégalité si et seulement si la suite $(x_n)_{n\in\mathbb{N}}$ des projections orthogonales de x sur les sous-espaces $\mathbb{E}_n = \operatorname{vect}(e_i)_{0\leqslant i\leqslant n}$

$$x_n = \sum_{i=0}^{n} \langle e_i, x\rangle\, e_i$$

converge dans l'espace normé $(\mathbb{E}, \|\ \|)$ vers x, ce qui revient exactement à dire que le vecteur x appartient à l'adhérence du sous-espace $\operatorname{vect}(e_n)_{n\in\mathbb{N}}$.

Nous verrons notamment une application de ce résultat dans le chapitre sur les séries de Fourier (égalité de Parseval Bessel).

15-2 Espaces hermitiens

Un espace vectoriel hermitien est un espace préhilbertien complexe de dimension finie. Il résulte immédiatement des études précédentes que, si $\mathbb{F}$ est un sous-espace d'un espace hermitien de dimension n $(\mathbb{E}_n, \langle\ ,\ \rangle)$, on a

$$\mathbb{E}_n = \mathbb{F} \oplus \mathbb{F}^{\perp}$$

EXERCICE 15-2.1 Si $(\mathbb{E}_n, \langle\ ,\ \rangle)$ est un espace hermitien, montrer que l'application

$$\mathbb{E}_n \to \mathbb{E}_n^* \quad x \mapsto \langle x, \bullet\rangle$$

est un isomorphisme **semi-linéaire** de $\mathbb{E}_n$ vers son dual.

15-2.1 Calculs dans une base

15-2.1.1 Adjointe d'une matrice

DÉFINITION 15-2.2 *Si $A \in \mathcal{M}_{p,n}(\mathbb{C})$, on appelle matrice adjointe de A la matrice*

$$A^* = {}^t\overline{A} \in \mathcal{M}_{n,p}(\mathbb{C})$$

transposée de la conjuguée de A.

PROPOSITION 15-2.3 L'application $A \mapsto A^*$ est un isomorphisme semi-linéaire de $\mathcal{M}_{p,n}(\mathbb{C})$ dans $\mathcal{M}_{n,p}(\mathbb{C})$. Pour $A \in \mathcal{M}_{p,n}(\mathbb{C})$ et $B \in \mathcal{M}_{m,p}(\mathbb{C})$, on a

$$(A^*)^* = A \textbf{ et } (BA)^* = A^*B^*$$

En particulier, lorsqu'on travaille avec des matrices carrées, l'application définie sur $\mathcal{M}_n(\mathbb{C})$ par $M \mapsto M^*$ est un automorphisme involutif de $\mathcal{M}_n(\mathbb{C})$ considéré comme $\mathbb{R}$-espace vectoriel. Les sous-espaces propres de cet automorphisme sont formés des matrices hermitiennes et antihermitiennes :

DÉFINITION 15-2.4 *Une matrice $M \in \mathcal{M}_n(\mathbb{C})$ est hermitienne si et seulement si $M = M^*$, elle est antihermitienne si $M = -M^*$. Les ensembles*

$$\mathcal{H}_n(\mathbb{C}) = \{M \in \mathcal{M}_n(\mathbb{C}) \mid M = M^*\} \text{ et } \mathcal{AH}_n(\mathbb{C}) = \{M \in \mathcal{M}_n(\mathbb{C}) \mid M = -M^*\}$$

sont deux sous-espaces du $\mathbb{R}$-espace vectoriel $\mathcal{M}_n(\mathbb{C})$ et la décomposition

$$\forall M \in \mathcal{M}_n(\mathbb{C}) \quad M = \frac{1}{2}(M + M^*) + \frac{1}{2}(M - M^*)$$

montre que

$$\mathcal{M}_n(\mathbb{C}) = \mathcal{H}_n(\mathbb{C}) \oplus_{\mathbb{R}} \mathcal{AH}_n(\mathbb{C})$$

REMARQUE 15-2.5 Les $\mathbb{R}$-ev $\mathcal{H}_n(\mathbb{C})$ et $\mathcal{AH}_n(\mathbb{C})$ sont isomorphes car il est clair que

$$\forall M \in \mathcal{M}_n(\mathbb{C}) \quad M \in \mathcal{H}_n(\mathbb{C}) \Leftrightarrow iM \in \mathcal{AH}_n(\mathbb{C})$$

Comme la dimension (réelle) de $\mathcal{M}_n(\mathbb{C})$ est égale à $2n^2$, il en résulte que

$$\dim_{\mathbb{R}} \mathcal{H}_n(\mathbb{C}) = \dim_{\mathbb{R}} \mathcal{AH}_n(\mathbb{C}) = n^2$$

Une matrice $A = (a_{ij})$ hermitienne a ses coefficients diagonaux réels, et les coefficients a_{ij} et a_{ji} symétriques par rapport à la diagonale principale sont conjugués.

PROPOSITION 15-2.6 Si $A \in \mathcal{M}_n(\mathbb{C})$, on a

$$\det A^* = \overline{\det A} \quad \textbf{et} \quad \chi_{A^*} = \overline{\chi_A}$$

En particulier, les valeurs propres de A^* sont les conjuguées des valeurs propres de A, avec les mêmes multiplicités.

Démonstration : On a en effet, pour $\lambda \in \mathbb{C}$

$$\begin{aligned}\chi_{A^*}(\lambda) &= \det(A^* - \lambda I_n) \\ &= \det\left({}^t\left(\overline{A - \overline{\lambda} I_n}\right)\right) = \overline{\det(A - \overline{\lambda} I_n)} = \overline{\chi_A(\overline{\lambda})}\end{aligned}$$

Les polynômes caractéristiques de A et A^* sont donc conjugués. ■

15-2.1.2 Expression du produit scalaire

Nous avons vu que, dans une base orthonormale d'un espace hermitien $(\mathbb{E}_n, \langle\, ,\, \rangle)$, l'expression du produit scalaire est très simple. Qu'en est-il dans une base quelconque ?

DÉFINITION 15-2.7 *Si $\mathcal{B} = (u_i)_{1 \leqslant i \leqslant n}$ est une base de $\mathbb{E}_n$, on appelle matrice du produit scalaire dans $\mathcal{B}$ la matrice*

$$A = (\langle u_i, u_j \rangle)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{H}_n(\mathbb{C})$$

C'est en effet une matrice hermitienne puisque $\langle u_j, u_i \rangle = \overline{\langle u_i, u_j \rangle}$.

PROPOSITION 15-2.8 Si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ est la matrice du produit scalaire dans la base $\mathcal{B} = (u_i)_{1 \leqslant i \leqslant n}$, pour $x, y \in \mathbb{E}_n$ représentés par les vecteurs colonnes X et Y dans $\mathcal{B}$

$$\langle x, y \rangle = {}^t\overline{X} A Y = X^* A Y$$

avec en particulier

$$\|x\|^2 = \langle x, x \rangle = X^* A X$$

Démonstration : On a en effet, par sesquilinéarité

$$\langle x, y \rangle = \left\langle \sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j \right\rangle = \sum_{i=1}^{n} \overline{x_i} \left(\sum_{j=1}^{n} \langle u_i, u_j \rangle y_j \right)$$

qui est exactement le produit de la ligne X^* par la colonne AY. ■

La matrice $A = (\langle u_i, u_j \rangle) \in \mathcal{H}_n(\mathbb{C})$ vérifie donc

$$\forall X \neq 0 \quad X^* A X > 0$$

ce qu'on traduit en disant que A est une matrice **hermitienne définie positive.** On notera $A \in \mathcal{H}_n^{*+}(\mathbb{C})$.

REMARQUE 15-2.9 Réciproquement, si M est une matrice hermitienne définie positive, l'expression

$$(X|Y) = X^* M Y$$

définit sur $\mathbb{C}^n \times \mathbb{C}^n$ (ou sur un espace complexe de dimension n en travaillant dans une base particulière) une application linéaire à droite et semi-linéaire à gauche, telle que

$$\forall X, Y \in \mathbb{C}^n \quad \overline{(X|Y)} = (X^* M Y)^* = Y^* M^* X = (Y|X)$$

puisque M est hermitienne, et qui vérifie

$$\forall X \in \mathbb{C}^n \quad X \neq 0 \Rightarrow (X|X) > 0$$

Il s'agit donc d'un produit scalaire sur $\mathbb{C}^n$, dont la matrice dans la base canonique est égale à M, comme on le vérifie facilement.

Lors d'un changement de base, la matrice du produit scalaire se transforme comme dans le cas réel, sans oublier l'importance de la conjugaison :

PROPOSITION 15-2.10 Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases d'un espace hermitien $(\mathbb{E}_n, \langle\ ,\ \rangle)$, les matrices A et A' du produit scalaire dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$ sont liées par

$$A' = P^* A P$$

où P est la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$.

Démonstration : Si C_i et C_j sont les colonnes d'indices i et j de la matrice P, le coefficient a'_{ij} de la matrice A' étant égal au produit scalaire $\langle u'_i, u'_j \rangle$ des vecteurs de la base $\mathcal{B}'$, on a

$$a'_{ij} = C_i^* A C_j$$

puisque nous connaissons u'_i et u'_j par leurs coordonnées dans $\mathcal{B}$. Lorsqu'on fait varier i et j, on obtient $A' = P^* A P$. ■

L'existence de bases orthonormales et le procédé d'orthonormalisation de Schmidt peuvent se traduire matriciellement, avec des démonstrations analogues à celles vues dans le cadre des espaces euclidiens :

PROPOSITION 15-2.11 Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est hermitienne définie positive si et seulement s'il existe une matrice $P \in \mathcal{GL}_n(\mathbb{C})$ telle que

$$A = P^* P$$

Plus précisément, si $A \in \mathcal{H}_n^{*+}(\mathbb{C})$, il existe une unique matrice triangulaire supérieure T à coefficients diagonaux réels strictement positifs telle que

$$A = T^* T$$

15-2.2 Adjoint d'un endomorphisme

THÉORÈME 15-2.12 (ET DÉFINITION) Si u est un endomorphisme d'un espace hermitien $(\mathbb{E}_n, \langle\ ,\ \rangle)$, il existe un unique endomorphisme de $\mathbb{E}_n$ appelé adjoint de u et noté u^* tel que

$$\forall\, x, y \in \mathbb{E}_n \quad \langle u(x), y \rangle = \langle x, u^*(y) \rangle$$

Démonstration : Comme dans le cas réel, on peut définir l'endomorphisme u^* par dualité. En travaillant dans une base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$, on peut aussi remarquer que, si u^* existe, on a

$$\forall\, x \in \mathbb{E}_n \quad u^*(x) = \sum_{i=1}^{n} \langle e_i, u^*(x) \rangle\, e_i$$

Ceci donne une expression de u^*

$$\forall x \in \mathbb{E}_n \quad u^*(x) = \sum_{i=1}^{n} \langle u(e_i), x \rangle e_i$$

On vérifie facilement qu'on définit ainsi un endomorphisme qui répond à la question (et qui ne dépend pas de la base orthonormale choisie, puisqu'il est unique à vérifier la propriété de définition de l'adjoint). ■

Les propriétés suivantes s'obtiennent comme dans le cas réel :

1. L'application $u \mapsto u^*$ est **semi-linéaire** et involutive de $\mathcal{L}(\mathbb{E}_n)$ dans lui-même : $(u^*)^* = u$ et
$$\forall u, v \in \mathcal{L}(\mathbb{E}_n) \quad \forall \lambda \in \mathbb{C} \quad (u + \lambda v)^* = u^* + \overline{\lambda} v^*$$
2. Pour u et $v \in \mathcal{L}(\mathbb{E}_n)$, on a
$$(uv)^* = v^* u^*$$
3. Si $u \in \mathcal{L}(\mathbb{E}_n)$, on a
$$\ker u^* = (\operatorname{Im} u)^{\perp} \quad \text{et} \quad \operatorname{Im} u^* = (\ker u)^{\perp}$$
4. Si u est un endomorphisme d'un espace hermitien $\mathbb{E}_n$, u et u^* ont même rang.
5. Il en résulte que
$$u \in \mathcal{GL}(\mathbb{E}_n) \Leftrightarrow u^* \in \mathcal{GL}(\mathbb{E}_n)$$
et on a alors
$$\left(u^{-1}\right)^* = (u^*)^{-1}$$
6. Si $\mathbb{F}$ est un sous-espace de $\mathbb{E}$, $\mathbb{F}$ est u-stable $\Leftrightarrow$ $\mathbb{F}^{\perp}$ est u^*-stable
7. En particulier, la recherche des hyperplans stables par un endomorphisme u se ramène à celle des droites vectorielles stables par u^*, c'est à dire à la recherche des vecteurs propres de u^*.
8. Si $\mathcal{B}$ est une base **orthonormale** de $\mathbb{E}_n$
$$M(u^*, \mathcal{B}) = (M(u, \mathcal{B}))^*$$
9. La propriété qui précède et la proposition **15-2.**6 montrent que
$$\det u^* = \overline{\det u} \quad \text{et} \quad \chi_{u^*} = \overline{\chi_u}$$
En particulier, les valeurs propres de u^* sont les conjuguées des valeurs propres de u, avec les mêmes multiplicités. Rappelons qu'une droite vectorielle propre pour u^* est orthogonale à un hyperplan stable par u.

EXERCICE 15-2.13 Montrer que, pour λ valeur propre de u, on a

$$\dim \ker (u - \lambda id_{\mathbb{E}_n}) = \dim \ker (u^* - \overline{\lambda} id_{\mathbb{E}_n})$$

15-2.3 Automorphismes et matrices unitaires

15-2.3.1 Automorphismes unitaires

On étudie ici les endomorphismes conservant la norme (correspondant aux automorphismes orthogonaux d'un espace euclidien).

THÉORÈME 15-2.14 Soit u un endomorphisme d'un espace hermitien $\mathbb{E}_n$. Les propriétés suivantes sont équivalentes :

i) u conserve la norme

$$\forall x \in \mathbb{E}_n \quad \|u(x)\| = \|x\|$$

ii) u conserve le produit scalaire.

iii) u est un automorphisme de $\mathbb{E}_n$ et $u^{-1} = u^*$

$$uu^* = u^*u = id_{\mathbb{E}_n}$$

iv) La matrice M de u dans une base orthonormale de $\mathbb{E}_n$ vérifie

$$M^*M = I_n$$

v) Il existe une base orthonormale de $\mathbb{E}_n$ transformée par u en une base orthonormale.

vi) u transforme toute base orthonormale de $\mathbb{E}_n$ en une base orthonormale.

La démonstration se fait comme dans le cas réel, avec utilisation de la formule de polarisation vue à la section **15-1.2.2**.

DÉFINITION 15-2.15 *Un automorphisme d'un espace hermitien qui conserve la norme est appelé automorphisme unitaire. On vérifie aisément que*

$$\mathcal{U}(\mathbb{E}_n) = \{u \in \mathcal{GL}(\mathbb{E}_n) \mid \forall x \in \mathbb{E}_n \quad \|u(x)\| = \|x\|\}$$

est un sous groupe de $(\mathcal{GL}(\mathbb{E}_n), \circ)$, appelé groupe unitaire de l'espace hermitien $\mathbb{E}_n$.

PROPOSITION 15-2.16 L'application déterminant induit un morphisme surjectif du groupe $(\mathcal{U}(\mathbb{E}_n), \circ)$ sur le groupe multiplicatif $(\mathbb{U}, \times)$ des nombres complexes de module 1. Son noyau

$$\mathcal{SU}(\mathbb{E}_n) = \{u \in \mathcal{U}(\mathbb{E}_n) \mid \det u = 1\}$$

est un sous-groupe de $(\mathcal{U}(\mathbb{E}_n), \circ)$, appelé groupe spécial unitaire de l'espace hermitien $\mathbb{E}_n$.

Démonstration : Si $u \in \mathcal{U}(\mathbb{E}_n)$, l'égalité $u^*u = id_{\mathbb{E}_n}$ entraîne

$$\det(u^*u) = 1 = \det u^* \det u = |\det u|^2$$

ce qui donne bien $\det u \in \mathbb{U}$. Comme pour $\theta \in \mathbb{R}$, on a

$$\det\left(e^{i\frac{\theta}{n}} id_{\mathbb{E}_n}\right) = e^{i\theta}$$

il est clair que le déterminant est surjectif de $\mathcal{U}(\mathbb{E}_n)$ vers $\mathbb{U}$. ■

REMARQUE 15-2.17 La situation est plus simple que dans le cas réel : on reconstitue facilement le groupe $\mathcal{U}(\mathbb{E}_n)$ à partir de $\mathcal{SU}(\mathbb{E}_n)$, puisque si $u \in \mathcal{U}(\mathbb{E}_n)$ avec $\det u = e^{i\theta}$ $(\theta \in \mathbb{R})$, on a $v = e^{-i\frac{\theta}{n}} u \in \mathcal{SU}(\mathbb{E}_n)$.

Comme $\mathbb{C}$ est algébriquement clos, les remarques qui suivent vont permettre d'obtenir facilement un résultat de diagonalisation concernant les automorphismes unitaires :

- Si $\mathbb{F}$ est un sous-espace vectoriel stable par un automorphisme unitaire u, son supplémentaire orthogonal est u-stable également, et les endomorphismes induits sont eux aussi des automorphismes unitaires

$$u|_{\mathbb{F}} \in \mathcal{U}(\mathbb{F}) \quad \text{et} \quad u|_{\mathbb{F}^\perp} \in \mathcal{U}(\mathbb{F}^\perp)$$

- Si $u \in \mathcal{U}(\mathbb{E}_n)$, les seules valeurs propres possibles de u sont des complexes de module 1

$$\mathrm{Sp}\, u \subset \mathbb{U}$$

et les sous-espaces propres pour des valeurs propres distinctes sont orthogonaux deux à deux :

En effet, si λ est une valeur propre de u et x est vecteur propre associé, l'égalité $\|u(x)\| = \|x\|$ impose $|\lambda| = 1$. Si λ et $\mu \in \mathbb{U}$ sont deux valeurs propres distinctes de u, on a

$$\forall x \in \ker(u - \lambda id_{\mathbb{E}_n}) \quad \forall y \in \ker(u - \mu id_{\mathbb{E}_n})$$
$$\langle x, y\rangle = \langle u(x), u(y)\rangle = \overline{\lambda}\mu \langle x, y\rangle = \frac{\mu}{\lambda}\langle x, y\rangle$$

ce qui entraîne $\langle x, y\rangle = 0$ et prouve l'orthogonalité des sous-espaces propres.

THÉORÈME 15-2.18 Si u est un automorphisme unitaire d'un espace hermitien $\mathbb{E}_n$, la somme directe orthogonale des sous-espaces propres de u est égale à $\mathbb{E}_n$. L'automorphisme u est donc diagonalisable. Un endomorphisme de $\mathbb{E}_n$ est unitaire si et seulement s'il existe une base orthonormale qui le diagonalise, avec des valeurs propres qui sont toutes de module 1.

Démonstration : Supposons $u \in \mathcal{U}(\mathbb{E}_n)$. Le spectre de u n'est pas vide. Le sous-espace

$$\mathbb{F} = \overset{\perp}{\bigoplus_{\lambda \in \mathrm{Sp}\, u}} \ker(u - \lambda id_{\mathbb{E}_n})$$

est u-stable, et donc $\mathbb{F}^\perp$ l'est également. L'endomorphisme v induit par u sur $\mathbb{F}^\perp$ a un spectre vide, donc $\mathbb{F}^\perp = \{0_{\mathbb{E}_n}\}$ et $\mathbb{F} = \mathbb{E}_n$, ce qui montre que u est diagonalisable. Une réunion de bases orthonormales des différents sous-espaces propres de u est une base orthonormale de $\mathbb{E}_n$ qui diagonalise u. Réciproquement, si $u \in \mathcal{L}(\mathbb{E}_n)$ est diagonalisable dans une base orthonormale avec valeurs propres de module 1, il transforme cette base orthonormale en une base qui est évidemment orthonormale, donc $u \in \mathcal{U}(\mathbb{E}_n)$. ■

15-2.3.2 Matrices unitaires

DÉFINITION 15-2.19 *Une matrice $M \in \mathcal{M}_n(\mathbb{C})$ est unitaire si et seulement si*

$$M^* M = I_n$$

DÉFINITION 15-2.20 *L'ensemble $\mathcal{U}_n(\mathbb{C})$ est un sous-groupe de $(\mathcal{GL}_n(\mathbb{C}), \times)$, isomorphe au groupe unitaire de l'espace $\mathbb{C}^n$ muni de sa structure hermitienne canonique. Le groupe $(\mathcal{U}_n(\mathbb{C}), \times)$ est appelé groupe unitaire d'ordre n.*

Si $M \in \mathcal{U}_n(\mathbb{C})$, on a

$$|\det M| = 1$$

et l'ensemble des matrices unitaires de déterminant égal à 1

$$\mathcal{SU}_n(\mathbb{C}) = \{M \in \mathcal{U}_n(\mathbb{C}) \mid \det M = 1\}$$

est un sous-groupe de $(\mathcal{U}_n(\mathbb{C}), \times)$, appelé groupe spécial unitaire d'ordre n.

Comme dans le cas réel, on a

PROPOSITION 15-2.21 Une matrice $M \in \mathcal{M}_n(\mathbb{C})$ est unitaire si et seulement si ses colonnes forment une base orthonormale de $\mathbb{C}^n$ muni de sa structure euclidienne canonique. Cette propriété est aussi vérifiée par les vecteurs lignes de M.

EXERCICE 15-2.22 Montrer que l'ensemble $\mathcal{SU}_2(\mathbb{C})$ est exactement l'ensemble des matrices de la forme

$$M = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$$

où a et b sont des complexes vérifiant $|a|^2 + |b|^2 = 1$. Montrer que, contrairement au groupe $(\mathcal{O}_2^+(\mathbb{R}), \times)$, $(\mathcal{SU}_2(\mathbb{C}), \times)$ n'est pas commutatif.

On peut aussi considérer les matrices unitaires comme matrices de passage entre bases orthonormales :

PROPOSITION 15-2.23 Soient $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base orthonormale d'un espace hermitien $(\mathbb{E}_n, (\,,\,))$ et $\mathcal{B}' = (e'_i)_{1 \leqslant i \leqslant n}$ une autre base, $P = \mathcal{M}_{\mathcal{B}}(\mathcal{B}')$ étant la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. On a

$$\mathcal{B}' \text{ orthonormale} \Leftrightarrow P \in \mathcal{U}_n(\mathbb{C})$$

Les formules de changement de bases orthonormales dans un espace hermitien sont donc simples : si $x \in \mathbb{E}_n$ est représenté dans $\mathcal{B}$ par la colonne X et dans $\mathcal{B}'$ par X', on a

$$X = PX' \Leftrightarrow X' = P^* X$$

Le théorème **15-2**.18 donne immédiatement :

THÉORÈME 15-2.24 Toute matrice $M \in \mathcal{U}_n(\mathbb{C})$ est diagonalisable : elle est semblable à une matrice diagonale à coefficients diagonaux de module 1, et plus précisément

$$\exists P \in \mathcal{U}_n(\mathbb{C}) \quad \exists (\theta_k)_{1\leqslant k\leqslant n} \in \mathbb{R}^n \quad M = P^{-1}\operatorname{Diag}\left(e^{i\theta_1},\ldots,e^{i\theta_n}\right)P$$

Réciproquement, toute matrice $M \in \mathcal{M}_n(\mathbb{C})$ pouvant s'écrire sous cette forme est produit de trois matrices unitaires et est donc unitaire.

EXERCICE 15-2.25 $\mathcal{M}_n(\mathbb{C})$ étant muni d'une norme quelconque, montrer que $\mathcal{U}_n(\mathbb{C})$ est une partie compacte et connexe par arcs de $\mathcal{M}_n(\mathbb{C})$.

15-2.4 Endomorphismes auto-adjoints

15-2.4.1 Définition

DÉFINITION 15-2.26 *Un endomorphisme u d'un espace hermitien $(\mathbb{E}_n, \langle\,,\,\rangle)$ est dit* ***auto-adjoint (ou hermitien)*** *si et seulement si $u^* = u$, ce qui revient à dire que*

$$\forall x, y \in \mathbb{E}_n \quad \langle u(x), y\rangle = \langle x, u(y)\rangle$$

On note $\mathcal{H}(\mathbb{E}_n)$ l'ensemble des endomorphismes auto-adjoints de $\mathbb{E}_n$.

REMARQUE 15-2.27 On en déduit évidemment que

$$u \in \mathcal{H}(\mathbb{E}_n) \Rightarrow \forall x \in \mathbb{E}_n \quad \langle u(x), x\rangle \in \mathbb{R}$$

Si $\mathcal{B}$ est une base orthonormale de $\mathbb{E}_n$, on a donc

$$u \in \mathcal{H}(\mathbb{E}_n) \Leftrightarrow M(u^*, \mathcal{B}) = M(u, \mathcal{B})$$

ce qui signifie que la matrice de u dans $\mathcal{B}$ est hermitienne. On a donc

PROPOSITION 15-2.28 L'ensemble des endomorphismes auto-adjoints d'un espace hermitien de dimension n est un $\mathbb{R}$-ev isomorphe à $\mathcal{H}_n(\mathbb{C})$, et par conséquent

$$\dim_{\mathbb{R}}(\mathcal{H}(\mathbb{E}_n)) = n^2$$

15-2.4.2 Réduction des endomorphismes hermitiens

THÉORÈME 15-2.29 Si $u \in \mathcal{H}(\mathbb{E}_n)$, le spectre de u est réel. L'espace $\mathbb{E}_n$ est somme directe orthogonale des sous-espaces propres de u. Il existe une base orthonormale de $\mathbb{E}_n$ qui diagonalise u. Réciproquement, tout endomorphisme de $\mathbb{E}_n$ diagonalisable dans une base orthonormale avec spectre réel est auto-adjoint.

Démonstration : Si $u \in \mathcal{H}(\mathbb{E}_n)$ et λ est valeur propre de u, si x est un vecteur propre associé on a

$$\langle u(x), x\rangle = \overline{\lambda}\,\|x\|^2 = \langle x, u(x)\rangle = \lambda\,\|x\|^2$$

ce qui entraîne $\lambda \in \mathbb{R}$ et donc $\mathrm{Sp}\,u \subset \mathbb{R}$. De plus, si $\lambda \neq \mu$ sont deux valeurs propres distinctes de u, on a pour $x \in \ker(u - \lambda id_{\mathbb{E}_n})$ et $y \in \ker(u - \mu id_{\mathbb{E}_n})$

$$\langle u(x), y\rangle = \lambda\langle x, y\rangle = \langle x, u(y)\rangle = \mu\langle x, y\rangle$$

(on a utilisé le fait que $\lambda \in \mathbb{R}$), ce qui montre que

$$\ker(u - \lambda id_{\mathbb{E}_n}) \perp \ker(u - \mu id_{\mathbb{E}_n})$$

On conclut ensuite exactement comme dans la démonstration du théorème **15-2**.18, en utilisant le fait que l'orthogonal d'un sous-espace u-stable est u^*-stable, donc ici stable par u. Pour la réciproque, un endomorphisme dont la matrice dans une base orthonormale est diagonale réelle possède dans cette base une matrice hermitienne, et est donc auto-adjoint. ■

Ce théorème a la traduction matricielle suivante, obtenue en travaillant avec les endomorphismes de $\mathbb{C}^n$ (hermitien canonique) associés canoniquement aux matrices considérées.

COROLLAIRE 15-2.30 Si A est une matrice hermitienne, son spectre est réel. Elle est diagonalisable avec matrice de passage unitaire :

$$\exists D \text{ diagonale réelle} \quad \exists U \in \mathcal{U}_n(\mathbb{C}) \quad A = UDU^{-1} = UDU^*$$

Cette propriété caractérise les matrices hermitiennes.

Le polynôme caractéristique d'une matrice hermitienne[1] est donc scindé dans $\mathbb{R}[X]$. On retrouve en particulier ce résultat pour les matrices symétriques réelles (puisque $\mathcal{S}_n(\mathbb{R}) \subset \mathcal{H}_n(\mathbb{C})$). Rappelons que ceci a été un point crucial dans la démonstration du théorème de réduction des endomorphismes symétriques donnée au chapitre précédent.

COROLLAIRE 15-2.31 Toute matrice antisymétrique réelle est diagonalisable sur $\mathbb{C}$, avec des valeurs propres qui sont imaginaires pures et une matrice de passage unitaire.

Démonstration : Si $A \in \mathcal{A}_n(\mathbb{R})$, on a $iA \in \mathcal{H}_n(\mathbb{C})$. ■

EXERCICE 15-2.32 Si u est un endomorphisme d'un espace hermitien $\mathbb{E}_n$, il existe une base orthonormale de $\mathbb{E}_n$ transformée par u en famille orthogonale (considérer l'endomorphisme hermitien u^*u).

[1] Ce polynôme est à coefficients réels puisque

$$A \in \mathcal{H}_n(\mathbb{C}) \Rightarrow \chi_{A^*} = \chi_A = \overline{\chi_A}$$

EXERCICE 15-2.33 Un endomorphisme d'un espace hermitien est dit normal s'il commute avec son adjoint. Montrer qu'un endomorphisme est normal si et seulement s'il est diagonalisable dans une base orthonormale.

EXERCICE 15-2.34 Soit $(u_i)_{i\in I}$ une famille d'endomorphismes auto-adjoints d'un espace hermitien $\mathbb{E}_n$. Montrer que les propriétés suivantes sont équivalentes :

i) $\forall\, i,j \in I \quad u_i u_j = u_j u_i$
ii) Il existe une base orthonormale diagonalisant tous les u_i
iii) $\exists\, u \in \mathcal{H}(\mathbb{E}_n) \quad \forall\, i \in I \quad \exists\, P_i \in \mathbb{R}[X] \quad u_i = P_i(u)$

15-2.4.3 Endomorphismes hermitiens positifs, définis positifs

Nous donnerons ici quelques résultats analogues à ceux obtenus dans le cas réel. Les démonstrations sont identiques.

Si $u \in \mathcal{H}(\mathbb{E}_n)$, nous avons vu que $\langle u(x), x\rangle \in \mathbb{R}$ pour tout x de $\mathbb{E}_n$. On dit que u est positif (nous noterons $u \in \mathcal{H}^+(\mathbb{E}_n)$) si

$$\forall\, x \in \mathbb{E}_n \quad \langle u(x), x\rangle \geqslant 0$$

De même, u est défini positif ($u \in \mathcal{H}^{*+}(\mathbb{E}_n)$) si

$$\forall x \in \mathbb{E}_n \quad x \neq 0_{\mathbb{E}_n} \Rightarrow \langle u(x), x\rangle > 0$$

A ces deux notions correspondent évidemment celles de matrices hermitiennes positives et définies positives :

$$A \in \mathcal{H}_n^+(\mathbb{C}) \Leftrightarrow \forall\, X \in \mathbb{C}^n \quad X^* A X \geqslant 0$$

$$A \in \mathcal{H}_n^{*+}(\mathbb{C}) \Leftrightarrow \forall\, X \neq 0_{\mathbb{C}^n} \quad X^* A X > 0$$

PROPOSITION 15-2.35 Si u est un endomorphisme auto-adjoint

$$u \in \mathcal{H}^+(\mathbb{E}_n) \Leftrightarrow \operatorname{Sp} u \subset \mathbb{R}^+$$

$$u \in \mathcal{H}^{*+}(\mathbb{E}_n) \Leftrightarrow \operatorname{Sp} u \subset \mathbb{R}^{*+}$$

PROPOSITION 15-2.36 Tout endomorphisme auto-adjoint positif possède une unique racine carrée hermitienne positive.

EXERCICE 15-2.37 Montrer que toute matrice $M \in \mathcal{GL}_n(\mathbb{C})$ admet une écriture unique sous la forme

$$A = UH$$

avec $U \in \mathcal{U}_n(\mathbb{C})$ et $H \in \mathcal{H}_n^{*+}(\mathbb{C})$. Montrer que l'application

$$\mathcal{GL}_n(\mathbb{C}) \to \mathcal{U}_n(\mathbb{C}) \times \mathcal{H}_n^{*+}(\mathbb{C}) \quad A \mapsto (U, H)$$

est continue. Indication : il s'agit de montrer que, si une suite de matrices inversibles $(A_p)_{p\in\mathbb{N}}$ se décomposant en $A_p = U_p H_p$ converge vers une matrice inversible $A = UH$, les suites (U_p) et (H_p) convergent (respectivement vers U et H). On pourra utiliser la compacité de $\mathcal{U}_n(\mathbb{C})$. (voir exercice **15-2.25**).

PROPOSITION 15-2.38 Si u est un endomorphisme d'un espace hermitien $(\mathbb{E}_n, (\ ,\ ))$, sa norme (en tant qu'endomorphisme continu) est donnée par

$$\|u\| = \sup_{\substack{\|x\| \leqslant 1 \\ \|y\| \leqslant 1}} |(u(x), y)|$$

On en déduit aisément que $\|u\| = \|u^*\|$.

PROPOSITION 15-2.39 Si u est un endomorphisme hermitien, on a également

$$\|u\| = \sup_{\|x\| \leqslant 1} |(u(x), x)| = \max\{|\lambda|,\ \lambda \in \operatorname{Sp} u\}$$

COROLLAIRE 15-2.40 Si $u \in \mathcal{L}(\mathbb{E}_n)$, la norme de u est la racine carrée de la plus grande valeur propre de u^*u.

15-3 Exercices

EXERCICE 15-3.1 Soit $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que $|\text{Trace}(A)| \leqslant \sqrt{n\text{Trace}(AA^*)}$. Quand y a-t-il égalité?

EXERCICE 15-3.2 Si A et B sont hermitiennes positives, montrer que

$$0 \leqslant \text{Trace}(AB) \leqslant \text{Trace}(A)\text{Trace}(B)$$

EXERCICE 15-3.3 Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne. Montrer que $\sum_{i,j} |a_{i,j}|^2$ est égal à la somme des carrés des valeurs propres de A.

EXERCICE 15-3.4 Soit $M \in \mathcal{M}_n(\mathbb{C})$.

1. Montrer qu'il existe une matrice de passage U unitaire trigonalisant M.
2. En déduire que, si le spectre de M est $\{\lambda_1 \cdots \lambda_n\}$ on a :

$$M^*M = MM^* \Leftrightarrow \sum_{i=0}^{n} |\lambda_i|^2 = \text{Trace}(M^*M)$$

3. En déduire que M est diagonalisable avec matrice de passage unitaire si et seulement si M et M^* commutent.

EXERCICE 15-3.5 Soient E hermitien, $u, v \in L(E)$. On note A le spectre de u^*u et B celui de v^*v. Montrer que toute valeur propre λ de $u \circ v$ vérifie :

$$\min A. \min B \leqslant |\lambda|^2 \leqslant \max A. \max B$$

Si H est hermitienne définie positive et α est un réel strictement positif, on pose

$$E_\alpha = \{A \text{ définies positives} \mid \det A \geqslant \alpha\}$$

Montrer que

$$\inf\{\text{Trace}(AH) \mid A \in E_\alpha\} = n(\alpha \det H)^{\frac{1}{n}}$$

Trouver de même pour H hermitienne positive

$$\max\{\text{Trace}(UH) \mid U \in U_n(\mathbb{C})\}$$

EXERCICE 15-3.6 Soit E un espace hermitien de dimension n et $a \in L(E)$. On note $\|.\|$ la norme de $L(E)$ subordonnée à la norme hermitienne et

$$\rho(a) = \max\{|\lambda|; \lambda \in \mathrm{sp}(A)\}$$

On suppose que $\|a\| \leqslant 1$.

1. Montrer que $\rho(a) \leqslant 1$. Montrer que $S = \{x \in E \mid \|a(x)\| = \|x\|\}$ est un sous-espace de E.

2. Montrer que $\rho(a) = 1$ ssi $\|a^n\| = 1$. (On pourra introduire un vecteur x_0 tel que $\|a(x_0)\| = \|x_0\|$ et montrer que le sev engendré par $\{x_0, a(x_0), \cdots, a^{n-1}(x_0)\}$ est stable par a et contenu dans S).

EXERCICE 15-3.7 Soit $A \in \mathcal{M}_n(\mathbb{R})$ antisymétrique. Montrer que $\det(I_n + A) \geqslant 1$. Montrer que, si $S \in \mathcal{M}_n(\mathbb{R})$ est symétrique positive, $\det(S + A) \geqslant \det S$.

EXERCICE 15-3.8 Soit $M \in \mathcal{GL}_n(\mathbb{C})$. Montrer que

$$\forall k \in \mathbb{N}^* \quad \exists ! : A \in \mathcal{GL}_n(\mathbb{C}) \quad A(A^*A)^k = M.$$

EXERCICE 15-3.9 Soit $A = \left(a_{i,j}\right) \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne positive. Montrer qu'il existe n vecteurs de $\mathbb{C}^n$ hermitien canonique $u_1, \cdots, u_n$ tels que

$$\forall i, j \quad a_{i,j} = < u_i, u_j >$$

EXERCICE 15-3.10 Soient E un espace hermitien de dimension n, u un opérateur auto-adjoint de E et la suite de ses valeurs propres $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ rangées par ordre décroissant. Soit $(a_i)_{1 \leqslant i \leqslant n}$ une base orthonormale de E telle que $\forall i \;\; u(a_i) = \lambda_i a_i$.

1. Montrer que

$$\lambda_p = \sup\left\{ < u(x), x > \mid \|x\| = 1 \text{ et } < x, a_k > = 0 \text{ pour } k = 1, \cdots, p-1\right\}$$

2. Pour $p \in \{1, \cdots, n\}$ et $\{z_1, \cdots, z_{p-1}\} \in E^{p-1}$ on pose

$$\begin{aligned}&\rho_u(z_1, \cdots, z_{p-1})\\ &\qquad = \sup\left\{ < u(x), x > \mid \|x\| = 1 \text{ et } < x, z_k > = 0 \text{ pour } k = 1, \cdots, p-1\right\}\end{aligned}$$

Montrer que $\rho_u(a_1, \cdots, a_{p-1}) \leqslant \rho_u(z_1, \cdots, z_{p-1})$.

3. On suppose de plus que $u = u' + u''$ où u' et u'' sont auto-adjoints. Montrer que $\lambda_{p+q-1} \leqslant \lambda'_p + \lambda''_q$ chaque fois que cela a un sens.

Chapitre 16

Séries de Fourier

16-1 Coefficients de Fourier

16-1.1 Espaces fonctionnels

16-1.1.1 Espace des fonctions 2π-périodiques

On considèrera dans ce chapitre l'espace $\mathbf{E}$ des fonctions $f : \mathbb{R} \to \mathbb{C}$, 2π-périodiques et continues par morceaux. Cela signifie qu'il existe une subdivision

$$0 = x_0 < x_1 < \cdots < x_n = 2\pi$$

de $[0, 2\pi]$ telle que

$$\forall i \in \{0, \ldots, n-1\} \quad f|_{]x_i, x_{i+1}[} \text{ se prolonge en une fonction } \mathcal{C}^0 \text{ sur } [x_i, x_{i+1}]$$

Par périodicité, il est clair que, sur tout segment, f n'admet qu'un nombre fini de discontinuités avec en tout point existence de limites à droite et à gauche.

On définit sur $\mathbf{E}$ un "pseudo-produit scalaire"

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t)\, dt$$

Il s'agit d'une forme hermitienne positive ($\forall f \in \mathbf{E} \quad \langle f, f \rangle \geqslant 0$) mais

$$\langle f, f \rangle = 0 \Leftrightarrow f \text{ est nulle sur } [0, 2\pi], \text{ sauf en un nombre fini de points.}$$

L'inégalité de Schwarz est toujours valable, mais

$$\|f\|_2 = \sqrt{(f,f)} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2 \, dt}$$

définit une **semi-norme** sur **E**. C'est la semi-norme associée à la convergence en moyenne quadratique. Si

$$C_{2\pi} = \{f \in \mathbb{E} \mid f \text{ continue sur } \mathbb{R}\}$$

la restriction de $(\, , \,)$ à $C_{2\pi}$ munit cet espace d'une structure d'espace préhilbertien complexe.

16-1.1.2 Polynômes trigonométriques

Dans $(C_{2\pi}, (\, , \,))$, on vérifie que la famille de fonctions $(e_n)_{n\in\mathbb{Z}}$ définie par

$$\forall t \in \mathbb{R} \quad e_n(t) = e^{int}$$

est **une famille orthonormale**. Elle est donc libre, et engendre un sous-espace (de fonctions indéfiniment dérivables)

$$T_{2\pi} = \text{vect}\,(e_n)_{n\in\mathbb{Z}}$$

appelé **espace vectoriel des polynômes trigonométriques 2π-périodiques**. Une fonction $f \in T_{2\pi}$ est donc caractérisée par la donnée d'une famille de complexes $(c_n)_{n\in\mathbb{Z}}$ **presque tous nuls** tels que

$$\forall t \in \mathbb{R} \quad f(t) = \sum_{k\in\mathbb{Z}} c_k e^{ikt}$$

Si f n'est pas nulle, en considérant $n = \max\{|k| \mid k \in \mathbb{Z} \text{ et } c_k \neq 0\}$, on pourra alors écrire

$$\forall t \in \mathbb{R} \quad f(t) = \sum_{k=-n}^{n} c_k e^{ikt}$$

et on dira que f est un polynôme trigonométrique de degré n. Pour $n \in \mathbb{N}$, l'espace vectoriel

$$T_{2\pi}^n = \text{vect}\,(e_k\,, \, -n \leqslant k \leqslant n)$$

est donc l'espace des polynômes trigonométriques de degré inférieur ou égal à n, et on a

$$T_{2\pi} = \bigcup_{n\in\mathbb{N}} T_{2\pi}^n$$

Surtout lorsqu'on travaille avec des fonctions réelles, il peut paraître plus simple de choisir d'autres bases des $T_{2\pi}^n$ (et donc de $T_{2\pi}$), en utilisant les relations

$$\forall k \in \mathbb{N} \quad \cos kt = \frac{e^{ikt} + e^{-ikt}}{2} \text{ et } \sin kt = \frac{e^{ikt} - e^{-ikt}}{2i}$$

En considérant les familles $\mathcal{C} = (t \mapsto \cos kt)_{k \geqslant 0}$ et $\mathcal{S} = (t \mapsto \sin kt)_{k>0}$, on vérifie aisément que $\mathcal{C} \cup \mathcal{S}$ est une base de $T_{2\pi}$, et que, pour $n \in \mathbb{N}$, la famille de fonctions $(t \mapsto \cos kt)_{n \geqslant k \geqslant 0} \cup (t \mapsto \sin kt)_{n \geqslant k > 0}$ est une base de $T_{2\pi}^n$: si f est un polynôme trigonométrique de degré inférieur à n, on pourra écrire

$$\forall t \in \mathbb{R} \quad f(t) = \sum_{k=-n}^{n} c_k e^{ikt} = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt)$$

les formules d'Euler montrant aisément que

$$\boxed{a_0 = c_0 \quad \text{et} \quad \forall k \geqslant 1 \begin{cases} a_k = c_k + c_{-k} \\ b_k = i\,(c_k - c_{-k}) \end{cases}}$$

ce qui peut s'écrire aussi

$$\boxed{\forall k \geqslant 1 \begin{cases} c_k = \dfrac{1}{2}(a_k - ib_k) \\ c_{-k} = \dfrac{1}{2}(a_k + ib_k) \end{cases}}$$

Ces nouvelles bases sont formées de fonctions réelles, forment des **systèmes orthogonaux** puisque

$$\forall k, l \quad \int_0^{2\pi} \cos kt \sin lt \, dt = 0$$

$$\text{et} \quad \forall k \neq l \quad \int_0^{2\pi} \cos kt \cos lt \, dt = \int_0^{2\pi} \sin kt \sin lt \, dt = 0$$

mais ces fonctions ne sont plus unitaires. On a en effet

$$\forall k \geqslant 1 \quad \frac{1}{2\pi} \int_0^{2\pi} \cos^2 kt \, dt = \frac{1}{2\pi} \int_0^{2\pi} \sin^2 kt \, dt = \frac{1}{2}$$

et donc, pour $k \geqslant 1$ $\|t \mapsto \cos kt\|_2 = \|t \mapsto \sin kt\|_2 = \dfrac{\sqrt{2}}{2}$. Elles sont donc d'un usage un peu moins commode pour tous les calculs de norme de convergence quadratique (voir plus loin la formule de Parseval).

Si f est un polynôme trigonométrique, on pourra donc écrire

$$\forall t \in \mathbb{R} \quad f(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ikt} = a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt)$$

(tous les coefficients étant nuls à partir d'un certain rang). S'agissant, pour la première écriture, de l'expression de f dans une base orthonormale, on a évidemment

$$\forall k \in \mathbb{Z} \quad c_k = (e_k, f) = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, e^{-ikt} dt$$

et, par les formules d'Euler

$$a_0 = c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,dt$$

$$\text{et} \quad \forall k \geqslant 1 \quad a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\,dt, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\,dt$$

16-1.2 Coefficients de Fourier d'une fonction 2π-périodique

Les intégrales précédentes, qui représentent les coordonnées d'un polynôme trigonométrique dans une base peuvent être évaluées pour n'importe quelle fonction continue par morceaux.

DÉFINITION 16-1.1 *Si f est continue par morceaux et 2π-périodique, on définit ses coefficients de Fourier (exponentiels) $(c_k(f))_{k\in\mathbb{Z}}$ par*

$$\forall k \in \mathbb{Z} \quad c_k(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,e^{-ikt}dt = \langle e_k, f\rangle$$

On utilisera aussi les coefficients de Fourier (trigonométriques) définis par

$$a_0(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,dt \quad \textit{et}$$

$$\forall k \geqslant 1 \quad a_k(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\,dt, \quad b_k(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\,dt$$

En tenant compte de la périodicité, on a aussi, pour tout a réel

$$c_k(f) = \frac{1}{2\pi}\int_a^{a+2\pi} f(t)\,e^{-ikt}dt$$

Les relations entre coefficients de Fourier exponentiels et trigonométriques sont les mêmes que dans la section précédente, puisqu'elles sont conséquences des formules d'Euler.

On est alors amené à considérer, pour $n \in \mathbb{N}$, les polynômes trigonométriques $(S_n^f)_{n\in\mathbb{N}}$ définis par

$$\forall t \in \mathbb{R} \quad S_n^f(t) = \sum_{k=-n}^{n} c_k(f)\,e^{ikt} = a_0(f) + \sum_{k=1}^{n}\left(a_k(f)\cos kt + b_k(f)\sin kt\right)$$

L'écriture $S_n^f = \sum_{k=-n}^{n} \langle e_k, f\rangle\, e_k$ montre que, dans le cas de fonctions continues, S_n^f **est projection orthogonale de f sur le sous-espace $T_{2\pi}^n$**. Dans le cas de fonctions continues par morceaux (où l'on ne travaille plus véritablement avec un produit scalaire), l'interprétation géométrique est analogue :

THÉORÈME 16-1.2 Si $f \in \mathbb{E}$ et $n \in \mathbb{N}$, le polynôme trigonométrique S_n^f est la meilleure approximation (au sens de la semi-norme de la convergence en moyenne quadratique) de f par un polynôme trigonométrique de degré inférieur ou égal à n.

Démonstration : C'est la même que dans le cadre strict des espaces préhilbertiens : si $s \in T_{2\pi}^n$, on a

$$\forall t \in \mathbb{R} \quad s(t) = \sum_{k=-n}^{n} \alpha_k e^{ikt} \text{ avec } \alpha_k \in \mathbb{C}$$

On écrit alors $f - s = \left(f - S_n^f\right) + \left(S_n^f - s\right)$, ce qui donne, puisque $f - S_n^f$ est orthogonal à $T_{2\pi}^n$

$$\|f - s\|_2^2 = \left\|f - S_n^f\right\|_2^2 + \sum_{k=-n}^{n} |c_k(f) - \alpha_k|^2$$

et donne évidemment

$$\|f - s\|_2 \geqslant \left\|f - S_n^f\right\|_2$$

avec égalité ssi pour tout $k \in [-n, n]$ on a $\alpha_k = c_k(f)$ soit $s = S_n^f$. ■

COROLLAIRE 16-1.3 (Inégalité de Bessel) Si $f \in \mathbb{E}$ et $n \in \mathbb{N}$, on a

$$\sum_{k=-n}^{n} |c_k(f)|^2 \leqslant \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 \, dt$$

En conséquence, la famille $\left(|c_n(f)|^2\right)_{n \in \mathbb{Z}}$ est sommable et

$$\sum_{n \in \mathbb{Z}} |c_n(f)|^2 \leqslant \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 \, dt$$

La série de terme général $\left(|a_k(f)|^2 + |b_k(f)|^2\right)$ est convergente et

$$|a_0(f)|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} \left(|a_n(f)|^2 + |b_n(f)|^2\right) \leqslant \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 \, dt$$

Démonstration : C'est une conséquence de l'égalité de Pythagore appliquée à l'écriture $f = S_n^f + \left(f - S_n^f\right)$. Comme pour $k \geqslant 1$,

$$|c_k(f)|^2 + |c_{-k}(f)|^2 = \frac{1}{2} \left(|a_k(f)|^2 + |b_k(f)|^2\right)$$

(égalité qui provient par exemple des calculs des normes de $t \mapsto \cos kt$ et $t \mapsto \sin kt$), on obtient aisément l'écriture du résultat en utilisant les coefficients trigonométriques. ■

REMARQUE 16-1.4 Nous verrons ultérieurement un résultat plus précis montrant qu'il y a en fait égalité entre la somme de la série et $\|f\|_2^2$ **(égalité de Parseval Bessel)**.

COROLLAIRE 16-1.5 Si $f \in \mathbb{E}$, on a

$$\lim_{n \to \pm\infty} c_n(f) = \lim_{n \to +\infty} a_n(f) = \lim_{n \to +\infty} b_n(f) = 0$$

16-1.3 Série de Fourier d'une fonction 2π-périodique

DÉFINITION 16-1.6 *Si $f \in \mathbb{E}$, on appelle série de Fourier de f la série de fonctions*

$$\sum_{n=0}^{+\infty} u_n(t) \quad \text{avec} \quad u_0(t) = a_0(f) \ \text{ et } \ \text{pour } n \geqslant 1$$

$$u_n(t) = a_n(f)\cos nt + b_n(f)\sin nt$$

La somme partielle d'ordre n de cette série est donc exactement

$$\sum_{k=0}^{n} u_k(t) = S_n^f(t) = \sum_{k=-n}^{n} c_k(f)\, e^{ikt} = a_0(f) + \sum_{k=1}^{n} \left(a_k(f)\cos kt + b_k(f)\sin kt\right)$$

Dans la suite de ce chapitre, on utilisera de deux façons différentes cette série de Fourier. Le plus souvent, **de manière formelle**. C'est-à-dire en ne considérant que les propriétés des suites de coefficients $(c_n(f))_{n\in\mathbb{Z}}$ (ou $(a_n(f))_{n\geqslant 0}$ et $(b_n(f))_{n>0}$), sans véritablement se poser la question de la convergence de la série. Nous étudierons également le problème de convergence (simple et uniforme) de la série et de l'égalité entre f et la somme de la série. Mais il faudra garder à l'esprit que **beaucoup de propriétés de f peuvent se lire sur la suite de ses coefficients de Fourier**. On a évidemment

PROPOSITION 16-1.7 Si $f \in T_{2\pi}$ est un polynôme trigonométrique 2π-périodique, on a

$$\forall t \in \mathbb{R} \quad f(t) = a_0(f) + \sum_{k=1}^{+\infty} \left(a_k(f)\cos kt + b_k(f)\sin kt\right)$$

(somme d'une série dont le terme général est nul à partir d'un certain rang).

16-1.4 Détermination pratique des coefficients de Fourier

On remarque d'abord que les coefficients de Fourier de f s'exprimant sous forme d'intégrales, ils ne changent pas si on modifie f en un nombre fini de points. Nous établissons dans cette section quelques résultats pratiques permettant parfois de simplifier le calcul des coefficients de Fourier.

PROPOSITION 16-1.8 L'application $\mathcal{F} : \mathbf{E} \to \mathbb{C}^{\mathbb{Z}}$ définie par

$$f \mapsto \mathcal{F}(f) = (c_n(f))_{n\in\mathbb{Z}}$$

est linéaire.

PROPOSITION 16-1.9 Si $f \in \mathbf{E}$ est réelle, on a

$$\forall k \in \mathbb{Z} \quad c_{-k}(f) = \overline{c_k(f)} \quad \textbf{et} \quad \forall k \quad a_k(f) \ \textbf{et} \ b_k(f) \in \mathbb{R}$$

PROPOSITION 16-1.10 Si $f \in \mathbb{E}$ est paire on a

$$\forall k > 0 \quad b_k(f) = 0 \quad \textbf{et}$$
$$a_k(f) = \frac{2}{\pi} \int_0^\pi f(t) \cos kt\, dt \quad \textbf{avec} \quad a_0(f) = \frac{1}{\pi} \int_0^\pi f(t)\, dt$$

Si f est impaire on a par contre

$$\forall k \geqslant 0 \quad a_k(f) = 0 \quad \textbf{et} \quad \forall k > 0 \quad b_k(f) = \frac{2}{\pi} \int_0^\pi f(t) \sin kt\, dt$$

Si f est paire, sa série de Fourier s'écrit donc

$$a_0(f) + \sum_{n=1}^{+\infty} a_n(f) \cos nt$$

et est composée uniquement de fonctions paires. Résultat analogue si f est impaire, indépendamment du problème de convergence de la série.

PROPOSITION 16-1.11 Si f est continue, 2π-périodique et $\mathcal{C}^1$ par morceaux, la fonction f', définie "presque partout" est prolongeable en une fonction continue par morceaux et 2π-périodique, prolongement que l'on note encore f', par abus d'écriture. On a alors

$$\forall k \in \mathbb{Z} \quad c_k(f') = ikc_k(f)$$
$$\forall k \geqslant 1 \quad a_k(f') = kb_k(f) \text{ et } b_k(f') = -ka_k(f)$$

Démonstration : Comme f est **continue et $\mathcal{C}^1$ par morceaux,** on peut calculer les coefficients de Fourier par intégration par parties :

$$c_k(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(t)\, e^{-ikt} dt$$
$$= \frac{1}{2\pi} \left(\left[f(t)\, e^{-ikt} \right]_0^{2\pi} + ik \int_0^{2\pi} f(t)\, e^{-ikt} dt \right) = ikc_k(f)$$

puisque $f(0) = f(2\pi)$. Le calcul est en tout point analogue pour les coefficients trigonométriques. ■

REMARQUE 16-1.12 Un moyen simple de retenir ces formules est de constater que, **formellement**, la série de Fourier de f' s'obtient en dérivant terme à terme celle de f :

$$\left(a_0(f) + \sum_{k=1}^{+\infty} \left(a_k(f) \cos kt + b_k(f) \sin kt \right) \right)' = \sum_{k=1}^{+\infty} \left(kb_k(f) \cos kt - ka_k(f) \sin kt \right)$$

ou aussi

$$\left(\sum_{k=-\infty}^{+\infty} c_k(f)\, e^{ikt} \right)' = \sum_{k=-\infty}^{+\infty} ikc_k(f)\, e^{ikt}$$

Il s'agit cependant d'un résultat obtenu par une simple intégration par parties, sans aucune utilisation du théorème de dérivation terme à terme (on n'est même pas assuré de la convergence des séries !).

COROLLAIRE 16-1.13 Si f est continue par morceaux, 2π-périodique avec

$$c_0(f) = 0$$

la fonction

$$F(x) = \int_0^x f(t)\, dt$$

est continue, $\mathcal{C}^1$ par morceaux et 2π-périodique, et ses coefficients de Fourier vérifient

$$\forall k \neq 0 \quad c_k(F) = \frac{c_k(f)}{ik}$$

Démonstration : La fonction F est évidemment continue et $\mathcal{C}^1$ par morceaux. De plus

$$\forall x \in \mathbb{R} \quad F(x+2\pi) - F(x) = \int_x^{x+2\pi} f(t)\, dt = \int_0^{2\pi} f(t)\, dt = 2\pi c_0(f)$$

et donc, la condition $c_0(f) = 0$ est nécessaire et suffisante pour que F soit 2π-périodique. Lorsque c'est le cas, on a $F'(x) = f(x)$ en tout point x de continuité de f, et il suffit alors d'appliquer le résultat précédent pour avoir

$$c_k(f) = ikc_k(F)$$

Lorsque $c_0(f) \neq 0$, c'est la fonction

$$x \mapsto \int_0^x [f(t) - c_0(f)]\, dt = F(x) - xc_0(f)$$

qui est périodique, et ses coefficients de Fourier peuvent s'obtenir facilement. ■

EXEMPLE 16-1.14 Soit f 2π-périodique vérifiant

$$f(x) = -1 \text{ sur }]-\pi, 0[\text{ et } f(x) = 1 \text{ sur }]0, \pi[$$

(on ne précise pas ses valeurs sur $\pi\mathbb{Z}$, ce qui n'a aucune importance pour le calcul des coefficients de Fourier). On vérifie alors que la fonction $F(x)$ vérifie

$$F(x) = |x| \text{ sur } [-\pi, \pi]$$

cette fonction étant ensuite prolongée par périodicité. Comme f est impaire, les $a_k(f)$ sont tous nuls et

$$b_k(f) = \frac{2}{\pi}\int_0^\pi \sin kt\, dt = \begin{cases} 0 \text{ si } k \text{ est pair} \\ \dfrac{4}{\pi k} \text{ si } k \text{ est impair} \end{cases}$$

La série de Fourier de f est donc

$$\frac{4}{\pi}\sum_{k=0}^{+\infty}\frac{\sin(2k+1)t}{2k+1}$$

Celle de F peut donc être obtenue par intégration terme à terme, sans oublier la "constante d'intégration"

$$c_0(F)=\frac{1}{\pi}\int_0^{\pi}t\,dt=\frac{\pi}{2}$$

Cette série de Fourier s'écrit donc

$$\frac{\pi}{2}-\frac{4}{\pi}\sum_{k=0}^{+\infty}\frac{\cos(2k+1)t}{(2k+1)^2}$$

La série de Fourier de F converge normalement sur $\mathbb{R}$. Celle de f converge (au moins simplement) sur $\mathbb{R}$, ce qu'on prouve par exemple par transformation d'Abel. Nous n'avons pas, pour le moment, d'information sur le lien entre les sommes de ces séries et les fonctions f et F.

EXERCICE 16-1.15 Soit f une fonction continue par morceaux et 2π-périodique et $x\in\mathbb{R}$ et $m\in\mathbb{Z}$. Exprimer à l'aide des coefficients $(c_k(f))_{k\in\mathbb{Z}}$ les coefficients de Fourier des fonctions

$$t\mapsto f(t-x)\ \text{ et }\ t\mapsto e^{imt}f(t)$$

Ici encore, une manipulation **formelle** de la série de Fourier de f permet de retenir ces résultats.

16-1.5 Coefficients de Fourier d'une série trigonométrique uniformément convergente

DÉFINITION 16-1.16 *On appelle série trigonométrique de période 2π, toute série de fonctions de la forme*

$$\alpha_0+\sum_{k=1}^{+\infty}(\alpha_k\cos kt+\beta_k\sin kt)$$

où $(\alpha_k)_{k\geqslant1}$ et $(\beta_k)_{k\geqslant1}$ sont deux suites complexes.

La somme partielle d'ordre n de cette série peut s'écrire

$$S_n(t)=\alpha_0+\sum_{k=1}^{n}(\alpha_k\cos kt+\beta_k\sin kt)=\sum_{k=-n}^{n}\gamma_k e^{ikt}$$

avec les relations entre les coefficients qui ont été vues au début de ce chapitre. Dans le cas de la convergence uniforme sur $\mathbb{R}$ de cette série, on peut faire le lien avec la notion de coefficients de Fourier :

THÉORÈME 16-1.17 Si la série trigonométique

$$S(t) = \alpha_0 + \sum_{k=1}^{+\infty} (\alpha_k \cos kt + \beta_k \sin kt)$$

converge uniformément sur $[0, 2\pi]$, sa somme S est une fonction continue sur $\mathbb{R}$, 2π-périodique, et ses coefficients de Fourier valent

$$a_0(S) = \alpha_0 \quad \text{et} \quad \forall k \geqslant 1 \quad a_k(S) = \alpha_k,\ b_k(S) = \beta_k$$

Démonstration : La continuité de S est conséquence de la convergence uniforme de la série. Si $n \in \mathbb{N}^*$, comme $t \mapsto \cos nt$ est bornée sur $[0, 2\pi]$, il y a convergence uniforme sur ce segment de la série

$$S(t) \cos nt = \alpha_0 \cos nt + \sum_{k=1}^{+\infty} (\alpha_k \cos kt + \beta_k \sin kt) \cos nt$$

On a donc

$$\begin{aligned} a_n(S) &= \frac{1}{\pi} \int_0^{2\pi} S(t) \cos nt\, dt \\ &= \frac{1}{\pi} \int_0^{2\pi} \alpha_0 \cos nt\, dt + \frac{1}{\pi} \int_0^{2\pi} \sum_{k=1}^{+\infty} (\alpha_k \cos kt + \beta_k \sin kt) \cos nt\, dt \end{aligned}$$

Les propriétés d'orthogonalité des fonctions de base de l'espace des polynômes trigonométriques montrent que toutes ces intégrales sont nulles, sauf celle correspondant à $k = n$, ce qui donne alors

$$a_n(S) = \frac{1}{\pi} \int_0^{2\pi} \alpha_n \cos^2 nt\, dt = \alpha_n$$

Le calcul des $b_n(S)$ et de $a_0(S)$ est analogue. ∎

REMARQUE 16-1.18 La convergence uniforme de la série trigonométrique est en particulier assurée lorsque

$$\sum_{n=1}^{+\infty} (|\alpha_n| + |\beta_n|) \quad \text{converge}$$

puisqu'il y a alors convergence normale. Par exemple, si on note

$$S(t) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{+\infty} \frac{\cos(2k+1)t}{(2k+1)^2}$$

on définit clairement une fonction S continue sur $\mathbb{R}$, 2π-périodique, qui a même série de Fourier que la fonction continue sur $\mathbb{R}$, 2π-périodique définie par

$$\forall t \in [-\pi, \pi] \quad F(t) = |t|$$

conformément au calcul effectué à la fin de la section précédente. Cette situation se rencontre dans un cadre plus général, comme le montre l'énoncé suivant :

THÉORÈME 16-1.19 Si f est une fonction continue, 2π-périodique et C^1 par morceaux, sa série de Fourier converge normalement sur $\mathbb{R}$.

Démonstration : Il y a bien convergence normale de cette série, puisque, pour $k \neq 0$, on a

$$\begin{aligned} |a_k(f)| + |b_k(f)| &= \frac{1}{k}\left(|b_k(f')| + |a_k(f')|\right) \\ &\leqslant \frac{1}{2}\left(\frac{1}{k^2} + |b_k(f')|^2\right) + \frac{1}{2}\left(\frac{1}{k^2} + |a_k(f')|^2\right) \end{aligned}$$

et parce que l'inégalité de Bessel (corollaire **16-1**.3) montre que la série majorante, de terme général

$$\frac{1}{k^2} + \frac{1}{2}\left(|a_k(f')|^2 + |b_k(f')|^2\right)$$

est une série convergente. ■

Nous verrons ultérieurement que la somme de cette série, fonction continue 2π-périodique qui a mêmes coefficients de Fourier (*cf.* théorème **16-1**.17) que f est en fait égale à f.

EXERCICE 16-1.20 Calculer, pour $n \in \mathbb{N}^*$, les intégrales

$$\int_{-\pi}^{\pi} \frac{\sin nt}{2+\cos t} dt \text{ et } \int_{-\pi}^{\pi} \frac{\cos nt}{2+\cos t} dt$$

Indication : il s'agit en fait de calculer des coefficients de Fourier de la fonction continue 2π-périodique

$$f(t) = \frac{1}{2+\cos t}$$

On peut pour cela trouver une série trigonométrique convergeant uniformément vers f. Il est plus simple de travailler avec des exponentielles complexes. On pourra notamment poser $z = e^{it}$ et écrire

$$f(t) = \frac{2z}{z^2+4z+1}$$

qu'on décomposera en éléments simples. L'un de ceux-ci peut être développé en série entière de la variable z (se souvenir que $|z| = 1$), l'autre en série entière de la variable z^{-1}.

REMARQUE 16-1.21 Il existe des séries trigonométriques convergeant pour tout $x \in \mathbb{R}$ qui ne sont pas séries de Fourier de fonctions continues par morceaux . C'est le cas notamment de

$$\sum_{n=1}^{+\infty} \frac{\sin nx}{\sqrt{n}}$$

qui converge en tout point de $\mathbb{R}$ (transformation d'Abel) et définit une fonction 2π-périodique continue sur $]0, 2\pi[$ (majoration fine du reste pour prouver la CVU de cette série sur tout segment inclus dans $]0, 2\pi[$, toujours par une transformation d'Abel). Cette série trigonométrique n'est pas une série de Fourier, puisque la suite de ses coefficients n'est pas de carré sommable (*cf.* corollaire **16-1**.3).

16-2 Convergence d'une série de Fourier

Soit $f \in \mathbb{E}$. On note

$$\forall t \in \mathbb{R} \quad S_n^f(t) = \sum_{k=-n}^{n} c_k(f)\, e^{ikt} = a_0(f) + \sum_{k=1}^{n} (a_k(f)\cos kt + b_k(f)\sin kt)$$

la somme partielle d'ordre n de sa série de Fourier.

16-2.1 Noyau de Dirichlet

On cherche ici une forme intégrale pour exprimer la somme partielle de la série de Fourier de f. Comme

$$c_k(f) = \frac{1}{2\pi}\int_0^{2\pi} f(u)\, e^{-iku} du$$

on a

$$S_n^f(t) = \sum_{k=-n}^{n} c_k(f)\, e^{ikt} = S_n^f(t) = \frac{1}{2\pi}\int_0^{2\pi}\left[\sum_{k=-n}^{n} e^{ik(t-u)}\right] f(u)\, du$$

La fonction

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx} = 1 + 2\sum_{k=1}^{n}\cos kx$$

est appelée noyau de Dirichlet. C'est une fonction $\mathcal{C}^\infty$, 2π-périodique paire, polynôme trigonométrique de degré n, qui vérifie

$$\frac{1}{2\pi}\int_0^{2\pi} D_n(t)\, dt = 1$$

Pour $t \notin 2\pi\mathbb{Z}$, on peut donner une autre expression de D_n :

$$D_n(t) = e^{-int}\left(\sum_{k=0}^{2n} e^{ikt}\right) = e^{-int}\frac{e^{i(2n+1)t} - 1}{e^{it} - 1}$$

$$= \frac{e^{i\left(n+\frac{1}{2}\right)t} - e^{i\left(n+\frac{1}{2}\right)t}}{e^{i\frac{t}{2}} - e^{-i\frac{t}{2}}} = \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}}$$

cette expression pouvant évidemment être prolongée par continuité sur $2\pi\mathbb{Z}$. On obtient finalement l'expression intégrale de la somme partielle de la série de Fourier de f :

$$S_n^f(t) = \frac{1}{2\pi}\int_0^{2\pi} D_n(t-u)\, f(u)\, du$$

$$= \frac{1}{2\pi}\int_{t-2\pi}^{t} D_n(u)\, f(t-u)\, du = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(u)\, f(t-u)\, du$$

par un changement de variable évident et en tenant compte de la 2π-périodicité des fonctions f et D_n. Pour étudier la convergence de la suite $\left(S_n^f(t)\right)_{n\in\mathbb{N}}$ vers une limite l, il s'agira de majorer la différence $S_n^f(t) - l$, qu'on écrira

$$\begin{aligned} S_n^f(t) - l &= \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(u)\, f(t-u)\, du - \frac{1}{2\pi}\int_{-\pi}^{\pi} l\, D_n(u)\, du \\ &= \frac{1}{2\pi}\int_0^{\pi} [f(t+u) + f(t-u) - 2l]\, D_n(u)\, du \end{aligned}$$

en tenant compte de la parité de D_n.

16-2.2 Théorème de Dirichlet

Le calcul préparatoire de la section précédent permet d'obtenir le résultat de convergence suivant :

THÉORÈME 16-2.1 Si f est continue par morceaux, 2π-périodique et possède en un point t_0 des demi-dérivées à droite et à gauche (en un sens généralisé)

$$f_g'(t_0) = \lim_{h\to 0^+} \frac{f(t_0-0) - f(t_0-h)}{h} \quad \textbf{et} \quad f_d'(t_0) = \lim_{h\to 0^+} \frac{f(t_0+h) - f(t_0+0)}{h}$$

alors la série de Fourier de f converge en t_0 et a pour somme

$$a_0(f) + \sum_{k=1}^{+\infty} \left(a_k(f)\cos kt_0 + b_k(f)\sin kt_0\right) = \frac{f(t_0+0) + f(t_0-0)}{2}$$

où $f(t_0+0)$ et $f(t_0-0)$ représentent respectivement les limites à droite et à gauche de f en t_0. Si de plus f est continue en t_0, la somme vaut évidemment $f(t_0)$.

Démonstration : On écrit, conformément au calcul effectué dans la section précédente,

$$\begin{aligned} S_n^f(t_0) - \frac{1}{2}[f(t_0+0) + f(t_0-0)] &= I_n + I_n' \quad \text{avec} \\ I_n = \frac{1}{2\pi}\int_0^{\pi} \frac{f(t_0+u) - f(t_0+0)}{\sin\frac{u}{2}} \sin\left(n+\frac{1}{2}\right)u\, du \quad &\text{et} \\ I_n' = \frac{1}{2\pi}\int_0^{\pi} \frac{f(t_0-u) - f(t_0-0)}{\sin\frac{u}{2}} \sin\left(n+\frac{1}{2}\right)u\, du & \end{aligned}$$

et on montre que chacune de ces deux intégrales tend vers 0. Pour I_n par exemple, on pourra écrire

$$\begin{aligned} I_n = \frac{1}{2\pi}\int_0^{\pi} &[f(t_0+u) - f(t_0+0)]\cos nu\, du \\ &+ \frac{1}{2\pi}\int_0^{\pi} \left[\frac{f(t_0+u) - f(t_0+0)}{\sin\frac{u}{2}}\cos\frac{u}{2}\right]\sin nu\, du \end{aligned}$$

et on peut interpréter ces intégrales comme coefficients de Fourier (à un facteur multiplicatif $\frac{1}{4}$ près). La première pour une fonction continue par morceaux 2π-périodique paire qui coïnciderait sur $]0,\pi[$, avec la fonction $u \mapsto f(t_0+u) - f(t_0+0)$, la seconde pour une fonction continue par morceaux 2π-périodique impaire qui coïnciderait sur $]0,\pi[$, avec la fonction $u \mapsto \dfrac{f(t_0+u) - f(t_0+0)}{\sin\frac{u}{2}} \cos\frac{u}{2}$, fonction effectivement prolongeable par continuité en 0 à droite à cause de l'hypothèse de dérivabilité faite sur f, puisque

$$\lim_{u \to 0^+} \frac{f(t_0+u) - f(t_0+0)}{\sin\frac{u}{2}} \cos\frac{u}{2} = 2f_d'(t_0)$$

Le corollaire **16-1**.5 montre que ces intégrales tendent vers 0. ■

L'exemple **16-1**.14 traité précédemment donne en particulier (figures **16.1** et **16.2**)

$$\forall x \in \ ]0,\pi[\quad 1 = \frac{4}{\pi} \sum_{n=0}^{+\infty} \frac{\sin(2n+1)x}{2n+1}$$

$$\forall x \in \ ]-\pi,\pi[\quad |x| = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=0}^{+\infty} \frac{\cos(2n+1)x}{(2n+1)^2}$$

pour $x = \frac{\pi}{2}$, la première série donne

$$\frac{\pi}{4} = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1}$$

résultat déjà obtenu à l'aide du développement en série entière de la fonction arctan. En faisant $x = 0$ dans la deuxième série, on obtient

$$\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$$

ce qui permet d'obtenir facilement

$$\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

REMARQUE 16-2.2 Arrêtons-nous un instant sur les hypothèses du théorème de Dirichlet. Elles portent essentiellement sur des propriétés locales de f, l'existence de dérivées à droite et à gauche en t_0, et la somme de la série de Fourier ne dépend elle aussi que du comportement de f au voisinage de t_0. Ceci est quand même assez surprenant, si l'on pense que les formules intégrales donnant les valeurs des coefficients de Fourier tiennent compte du comportement global

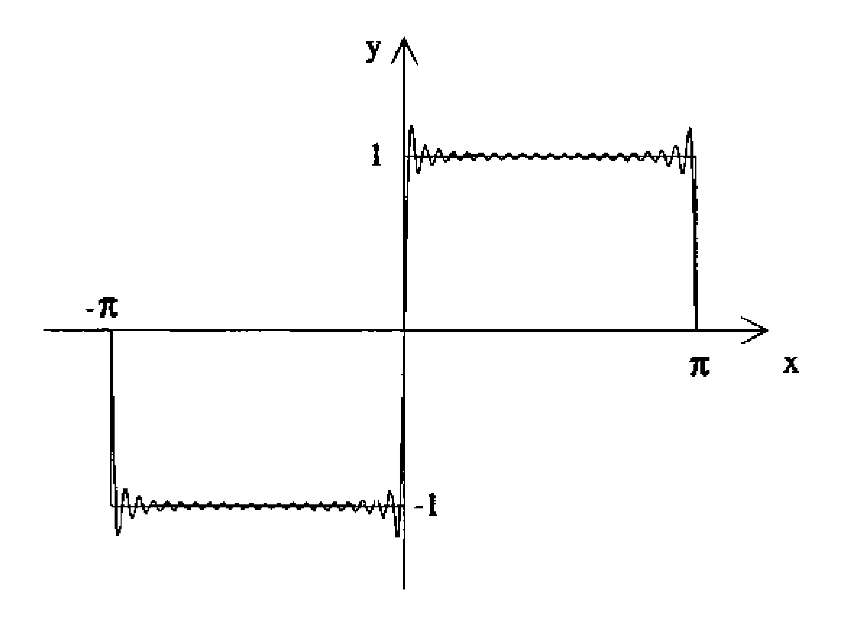


Figure **16.1** – *Approximation de la fonction "créneau" par une somme partielle de sa série de Fourier*

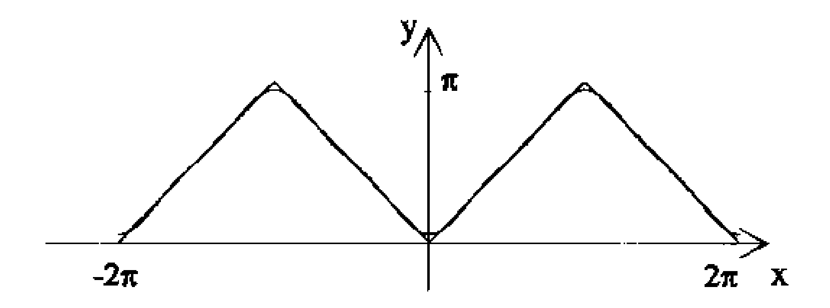


Figure **16.2** – *Approximation de la fonction "dent de scie" par une somme partielle de sa série de Fourier*

de f, sur tout un intervalle d'amplitude 2π. On s'aperçoit en conséquence qu'une fonction simple, disons $x \to x$ sur $]0,\pi[$ peut être représentée de différentes manières comme somme d'une série trigonométrique. En prolongeant cette fonction par parité et périodicité, nous avons obtenu plus haut

$$\forall x \in]0,\pi[\quad x = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=0}^{+\infty} \frac{\cos(2n+1)x}{(2n+1)^2}$$

En prolongeant cette fonction par imparité et périodicité, on obtiendrait une série de sinus :

$$\forall x \in]0,\pi[\quad x = 2\sum_{n=1}^{+\infty} (-1)^{n+1} \frac{\sin nx}{n}$$

Comme la somme des termes de rang impair vaut $\frac{\pi}{2}$ sur cet intervalle (c'est toujours l'exemple étudié précédemment), on pourrait aussi écrire

$$\forall x \in]0,\pi[\quad x = \frac{\pi}{2} + \sum_{n=1}^{+\infty} \frac{\sin 2nx}{n}$$

Il faut donc noter la différence très importante avec ce qui se passe pour les développements en série entière. On se gardera bien notamment d'identifier sans justification des coefficients de séries trigonométriques. Lorsqu'on voudra le faire, c'est souvent un résultat comme le théorème **16-1**.17 qu'on appliquera.

A l'usage, on utilisera plutôt les corollaires suivants, qu'on appellera également "théorèmes de Dirichlet" :

THÉORÈME 16-2.3 Si f est une fonction 2π-périodique et C^1 par morceaux, sa série de Fourier converge en tout point $t \in \mathbb{R}$ et sa somme vaut

$$\forall t \in \mathbb{R} \quad a_0(f) + \sum_{k=1}^{+\infty} (a_k(f)\cos kt + b_k(f)\sin kt) = \frac{f(t+0) + f(t-0)}{2}$$

Dans le cas d'une fonction supposée de plus continue, on aura, à l'aide du théorème **16-1**.19

THÉORÈME 16-2.4 Si f est une fonction <u>continue</u>, 2π-périodique et C^1 par morceaux, sa série de Fourier converge normalement sur $\mathbb{R}$ et

$$\forall t \in \mathbb{R} \quad a_0(f) + \sum_{k=1}^{+\infty} (a_k(f)\cos kt + b_k(f)\sin kt) = f(t)$$

On pourra lire quelquefois qu'une telle fonction f est "développable en série de Fourier" (c'est-à-dire que sa série de Fourier converge en tout point et a pour somme f), mais cette expression usuelle est malheureuse, car elle risque d'amener beaucoup de confusions avec la notion de développement en série entière, où l'on ne considère les fonctions que "localement".

16-2.3 Théorèmes de Weierstrass

Commençons par un premier résultat concernant les fonctions continues 2π-périodiques :

THÉORÈME 16-2.5 Si $f : \mathbb{R} \to \mathbb{C}$ est continue 2π-périodique, il existe une suite de polynômes trigonométriques 2π-périodiques qui converge uniformément vers f sur $\mathbb{R}$.

Démonstration : On caractérise ainsi les fonctions continues 2π-périodiques. Choisissons un réel $\varepsilon > 0$ arbitraire. Si f est comme dans l'énoncé du théorème, une subdivision à pas constant de $[0, 2\pi]$ par les points $\left(x_i = \frac{2i\pi}{p}\right)_{0 \leqslant i \leqslant p}$ permet par interpolation de définir une fonction g_p affine par morceaux continue sur $[0, 2\pi]$:

$$\forall x \in [x_i, x_{i+1}] \quad g_p(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i)$$

Comme f est 2π périodique, on a clairement

$$g_p(2\pi) = f(2\pi) = f(0) = g_p(0)$$

ce qui permet de prolonger par périodicité la fonction g_p en une fonction 2π-périodique continue et C^1 par morceaux. De plus, si p est assez grand, la continuité uniforme de f montre que

$$\|f - g_p\|_{\infty,[0,2\pi]} = \|f - g_p\|_{\infty,\mathbb{R}} \leqslant \frac{\varepsilon}{2}$$

(voir le théorème **8-2**.8). Fixant un tel p, la série de Fourier de g_p converge normalement vers g_p sur $\mathbb{R}$ (théorème **16-2**.4) et on peut alors trouver un indice n avec

$$\|S_n^{g_p} - g_p\|_{\infty,[0,2\pi]} \leqslant \frac{\varepsilon}{2}$$

On obtient alors

$$\|f - S_n^{g_p}\|_{\infty,[0,2\pi]} \leqslant \varepsilon$$

et donne une approximation uniforme de f par un polynôme trigonométrique. ■

Si f est à présent continue T-périodique ($T > 0$), la fonction $f_1(t) = f\left(\frac{T}{2\pi}t\right)$ est évidemment de période 2π. En lui appliquant le théorème précédent on obtient aisément :

COROLLAIRE 16-2.6 Si $f : \mathbb{R} \to \mathbb{C}$ est continue T-périodique, il existe une suite de polynômes trigonométriques T-périodiques qui converge uniformément vers f sur $\mathbb{R}$

Ici, un polynôme trigonométrique T-périodique est une fonction appartenant à l'espace vect $\left(t \mapsto e^{ik\frac{2\pi}{T}t}\right)_{k\in\mathbb{Z}}$, soit encore une fonction combinaison linéaire de la famille

$$\left(t \mapsto \cos k\frac{2\pi}{T}t\right)_{k\in\mathbb{N}} \cup \left(t \mapsto \sin k\frac{2\pi}{T}t\right)_{k\in\mathbb{N}^*}$$

On peut aussi sortir du cadre des fonctions périodiques pour obtenir le théorème d'approximation polynomiale annoncé à la section **8-2.3** :

THÉORÈME 16-2.7 Si $f : [a,b] \to \mathbb{R}$ ou $\mathbb{C}$ est une fonction continue, il existe une suite $(P_n)_{n\in\mathbb{N}}$ de polynômes de $\mathbb{R}[X]$ ou $\mathbb{C}[X]$ convergeant uniformément vers f sur $[a,b]$.

Démonstration : En considérant la fonction $g : [0,\pi] \to \mathbb{R}$ ou $\mathbb{C}$ définie par

$$g(t) = f\left(a + \frac{t}{\pi}(b-a)\right)$$

on se ramène à travailler sur $[0,\pi]$ ce qui est suffisant pour conclure dans le cas général, puisqu'en composant un polynôme avec une fonction affine, on retrouve un polynôme. On peut ensuite prolonger g par parité et 2π-périodicité pour obtenir une fonction à laquelle on peut appliquer le théorème précédent. On peut trouver un polynôme trigonométrique S vérifiant

$$\|g - S\|_{\infty,[0,\pi]} \leqslant \frac{\varepsilon}{2}$$

S est somme sur $\mathbb{R}$ d'une série entière de rayon de convergence infini (combinaison linéaire de fonctions exponentielles). Cette série entière convergeant uniformément sur $[0,\pi]$, on peut trouver une de ses sommes partielles S_n vérifiant

$$\|S - S_n\|_{\infty,[0,\pi]} \leqslant \frac{\varepsilon}{2}$$

ce qui donnera bien $\|g - S_n\|_{\infty,[0,\pi]} \leqslant \varepsilon$. ■

EXERCICE 16-2.8 Soit f une fonction continue $[a,b] \to \mathbb{C}$ telle que

$$\forall n \in \mathbb{N} \quad \int_a^b f(t)\, t^n\, dt = 0$$

Montrer que f est identiquement nulle.

16-2.4 Deux exercices

16-2.4.1 Développements eulériens

EXERCICE 16-2.9 Soit $z \in \mathbb{C} - \mathbb{Z}$. On considère la fonction continue et C^1 par morceaux définie par

$$|t| \leqslant \pi \Rightarrow f(t) = \cos(zt) = \frac{e^{izt} + e^{-izt}}{2}$$

et prolongée par 2π-périodicité. En déterminant la série de Fourier de f, montrer que

$$\forall t \in [-\pi, \pi] \quad \forall z \in \mathbb{C} - \mathbb{Z} \quad \cos(zt) = \sum_{n=-\infty}^{+\infty} \frac{(-1)^n z \sin(\pi z)}{\pi(z^2 - n^2)} e^{int}$$

(cette somme étant à prendre *a priori* comme limite des sommes partielles "symétriques" $\sum_{k=-n}^{k=n}$ pour $n \to +\infty$, mais on peut remarquer qu'en l'occurence, il s'agit aussi de la somme d'une famille sommable de complexes indexée par $\mathbb{Z}$).

En déduire que

$$\forall z \in \mathbb{C} - \mathbb{Z} \quad \pi \operatorname{cotan}(\pi z) = \sum_{n=-\infty}^{+\infty} \frac{z}{z^2 - n^2} = \frac{1}{z} + 2\sum_{n=1}^{+\infty} \frac{z}{z^2 - n^2}$$

$$\forall z \in \mathbb{C} - \mathbb{Z} \quad \frac{\pi}{\sin(\pi z)} = \frac{1}{z} + 2\sum_{n=1}^{+\infty} \frac{(-1)^n z}{z^2 - n^2}$$

En se limitant au cas réel avec $x \in]-1, 1[$, montrer que

$$\pi \operatorname{cotan}(\pi x) - \frac{1}{x} = 2\sum_{n=1}^{+\infty} \frac{x}{x^2 - n^2}$$

(quel sens donner au terme de gauche pour $x = 0$?). En déduire

$$x \in]-1, 1[\Rightarrow \sin(\pi x) = \pi x \prod_{n=1}^{+\infty} \left(1 - \frac{x^2}{n^2}\right)$$

Donner de même une écriture comme produit infini de $\operatorname{sh}(\pi x)$.

16-2.4.2 Théorème de Fejér

La démonstration précédente du théorème d'approximation d'une fonction continue périodique par des polynômes trigonométriques a le défaut de ne pas être très constructive. De plus, les hypothèses du théorème de Dirichlet font intervenir des propriétés de dérivabilité qui peuvent sembler superflues : qu'obtiendrait-on en supposant seulement les fonctions continues ?

On démontre que la série de Fourier d'une fonction continue converge "souvent", mais qu'il peut exister des points où cette série est divergente. Le procédé de Cesàro permet cependant d'obtenir alors une convergence, comme le montre le

THÉORÈME 16-2.10 Si f est continue par morceaux 2π-périodique, en tout point t_0 de continuité de f sa série de Fourier converge vers $f(t_0)$ en moyenne de Cesàro. Si f est de plus continue sur $\mathbb{R}$, la convergence est uniforme sur $\mathbb{R}$: les "sommes de Fejér"

$$\Sigma_n^f(t) = \frac{S_0^f(t) + \cdots + S_n^f(t)}{n+1}$$

forment alors une suite de polynômes trigonométriques qui converge uniformément vers f.

Démonstration : Indications : montrer que

$$\Sigma_n^f(t) = \frac{1}{2\pi}\int_0^{2\pi} f(t-u)\, F_n(u)\, du$$

où F_n est le noyau de Fejér, donné en fonction du noyau de Dirichlet par

$$F_n(u) = \frac{1}{n+1}\left[D_0(u) + \cdots + D_n(u)\right]$$

Montrer que l'on a

$$F_n(u) = \frac{\sin^2 \frac{n+1}{2} u}{(n+1)\sin^2 \frac{u}{2}}$$

expression prolongée par continuité sur $2\pi\mathbb{Z}$. Montrer que F_n est paire, vérifie

$$\frac{1}{2\pi}\int_0^{2\pi} F_n(u)\, du = 1$$

et que, pour $n \to +\infty$, F_n converge uniformément vers 0 sur tout intervalle $[\delta, \pi]$ avec $\delta > 0$. En déduire que

$$0 < \delta < \pi \Rightarrow \lim_{n\to+\infty} \frac{1}{2\pi}\int_{-\delta}^{+\delta} F_n(u)\, du = 1$$

Le fait que $F_n \geqslant 0$ va rendre la majoration de la différence

$$\Sigma_n^f(t_0) - f(t_0)$$

plus simple que celle de $S_n^f(t_0) - f(t_0)$ dans la démonstration du théorème de Dirichlet : écrire

$$\Sigma_n^f(t_0) - f(t_0) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \left[f(t_0 - u) - f(t_0)\right] F_n(u)\, du$$

Pour rendre cette différence inférieure à $\varepsilon > 0$ donné arbitrairement, on utilisera la continuité de f en t_0, ce qui amènera à découper l'intégrale précédente en morceaux. Dans le cas d'une fonction continue sur $\mathbb{R}$, on utilisera de plus un argument de continuité uniforme (voir section **12-6.3**). ■

Comme la convergence "classique" entraîne la convergence au sens de Cesàro, on déduit facilement du théorème précédent que, si la série de Fourier d'une fonction f continue périodique converge en un point t_0, sa somme est nécessairement égale à $f(t_0)$.

16-2.5 Convergence en moyenne quadratique

Soit f une fonction continue par morceaux de $[0, 2\pi] \to \mathbb{C}$. En modifiant f en 2π, (ce qui ne change rien pour f vis-à-vis du calcul intégral), on peut supposer $f(0) = f(2\pi)$. On peut alors prolonger f en une fonction 2π-périodique, et considérer notamment la série de Fourier de cette fonction. Nous l'appellerons, par abus de langage, série de Fourier de f.

THÉORÈME 16-2.11 Si f est une fonction continue par morceaux de $[0, 2\pi]$ dans $\mathbb{C}$, les sommes partielles de sa série de Fourier convergent en moyenne quadratique sur $[0, 2\pi]$:

$$\lim_{n\to+\infty} \int_0^{2\pi} \left|f(t) - S_n^f(t)\right|^2 dt = 0$$

Ce résultat peut s'écrire également

$$\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2 dt = |a_0(f)|^2 + \frac{1}{2}\sum_{n=1}^{+\infty}\left(|a_n(f)|^2 + |b_n(f)|^2\right) = \sum_{n\in\mathbb{Z}} |c_n(f)|^2$$

(égalité dite de Parseval-Bessel).

Démonstration : Si on note comme précédemment e_k la fonction $t \mapsto e^{ikt}$, on sait que

$$S_n^f = \sum_{k=-n}^{n} (e_i, f)\, e_i$$

est meilleure approximation en moyenne quadratique de f par un polynôme trigonométrique de degré inférieur ou égal à n. La relation de Pythagore donne par ailleurs

$$\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2 dt = \|f\|_2^2$$
$$= \|f - S_n^f\|_2^2 + \|S_n^f\|^2 = \|f - S_n^f\|_2^2 + \sum_{k=-n}^{n} |c_k(f)|^2$$

et montre que l'égalité de Parseval-Bessel traduit exactement la convergence en moyenne quadratique de S_n^f vers f. Soit $\varepsilon > 0$. Commençons par montrer qu'on peut trouver un polynôme trigonométrique T tel que

$$\|f - T\|_2 \leqslant \varepsilon$$

Pour ce faire, f ne présentant qu'un nombre fini de points de discontinuité sur $[0, 2\pi]$ on peut trouver une fonction continue g sur $[0, 2\pi]$ vérifiant $g(0) = g(2\pi)$ et $\|f - g\|_2 \leqslant \frac{\varepsilon}{2}$: si x_i est un point de discontinuité de f (intérieur à l'intervalle), il suffit de choisir $\alpha > 0$ assez petit et de remplacer sur $[x_i - \alpha, x_i + \alpha]$ la fonction f par une fonction

affine vérifiant $g(x_i \pm \alpha) = f(x_i \pm \alpha)$, et de faire de même sur $[0, \alpha]$ et $[2\pi - \alpha, 2\pi]$, avec les conditions $g(0) = g(2\pi) = 0$ par exemple. Le théorème de Weierstrass donne alors un polynôme trigonométrique T vérifiant $\|T - g\|_\infty \leqslant \frac{\varepsilon}{2}$, ce qui entraîne $\|T - g\|_2 \leqslant \frac{\varepsilon}{2}$ et, par inégalité triangulaire $\|f - T\|_2 \leqslant \varepsilon$. Si p est alors le degré de T, on aura, par propriété de meilleure approximation

$$\left\|f - S_p^f\right\|_2 \leqslant \|f - T\|_2 \leqslant \varepsilon$$

et comme la suite $n \mapsto \left\|f - S_n^f\right\|_2$ est décroissante

$$\forall n \geqslant p \quad \left\|f - S_n^f\right\|_2 \leqslant \varepsilon$$

ce qui prouve bien la convergence en moyenne quadratique. ■

Si le carré de la (semi)-norme peut s'exprimer à l'aide des coefficients de Fourier, il en est de même du (pseudo)-produit scalaire. On a notamment le

COROLLAIRE 16-2.12 Si f et g sont deux fonctions continues par morceaux $[0, 2\pi] \to \mathbb{C}$, on a

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \overline{f(t)}g(t)\,dt &= \sum_{n\in\mathbb{Z}} \overline{c_n(f)}c_n(g) \\ &= \overline{a_0(f)}a_0(g) + \frac{1}{2}\sum_{n=1}^{+\infty}\left(\overline{a_n(f)}a_n(g) + \overline{b_n(f)}b_n(g)\right)\end{aligned}$$

Démonstration : Remarquons tout d'abord que la famille de complexes $\left(\overline{c_n(f)}c_n(g)\right)_{n\in\mathbb{Z}}$ est sommable, à cause de la majoration classique

$$\left|\overline{c_n(f)}c_n(g)\right| \leqslant \frac{1}{2}\left(|c_n(f)|^2 + |c_n(g)|^2\right)$$

La formule pourrait être démontrée par "polarisation". On peut aussi utiliser la continuité du (pseudo)-produit scalaire, et remarquer que

$$\lim_{n\to+\infty} \left\langle S_n^f, S_n^g\right\rangle = \langle f, g\rangle$$

Comme

$$\left\langle S_n^f, S_n^g\right\rangle = \sum_{k=-n}^{n} \overline{c_k(f)}c_k(g)$$

(produit scalaire de deux polynômes trigonométriques), on obtient facilement le résultat. ■

EXERCICE 16-2.13 En reprenant la fonction $g(x) = |x|$ sur $[-\pi, \pi]$ prolongée par périodicité étudiée à l'exemple **16-1**.14, montrer que

$$\zeta(4) = \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

En intégrant la fonction $t \mapsto g(t) - \frac{\pi}{2}$, obtenir de même

$$\zeta(6) = \frac{\pi^6}{945}$$

On voit apparaître un procédé (ce n'est pas le plus habile) de calcul de proche en proche des valeurs de la fonction ζ de Riemann sur les entiers pairs.

EXERCICE 16-2.14 Soit f une fonction de classe C^1 sur $[0, \pi] \to \mathbb{R}$, vérifiant $f(0) = f(\pi) = 0$. Déterminer la plus petite constante (indépendante de f) telle que

$$\int_0^\pi f^2(t)\, dt \leqslant C \int_0^\pi [f'(t)]^2\, dt$$

et traiter les cas d'égalité.

16-2.6 Exercice : convolution des fonctions périodiques

Si f et g sont deux fonctions continues par morceaux et 2π-périodiques, on définit leur produit de convolution par

$$\forall x \in \mathbb{R} \quad f * g(x) = \frac{1}{2\pi} \int_0^{2\pi} f(x-t)\, g(t)\, dt$$

- Montrer que cette fonction est 2π-périodique, et que $f * g = g * f$.
- Exprimer, à l'aide des $(c_k(f))_{k \in \mathbb{Z}}$, les coefficients de Fourier de la fonction $t \mapsto \overline{f(x-t)}$.
- En déduire que

$$\forall x \in \mathbb{R} \quad f * g(x) = \sum_{k \in \mathbb{Z}} c_k(f)\, c_k(g)\, e^{ikx}$$

- Montrer que $f * g$ est continue sur $\mathbb{R}$, et vérifie

$$\forall k \in \mathbb{Z} \quad c_k(f * g) = c_k(f)\, c_k(g)$$

  En déduire que le produit de convolution est associatif dans l'ensemble des fonctions continues 2π-périodiques, distributif par rapport à l'addition, mais qu'il ne possède pas d'élément neutre.

On pourrait interpréter les calculs relatifs aux noyaux de Dirichlet et Fejér avec le produit de convolution : si f est continue par morceaux 2π-périodique, quels sont les coefficients de Fourier de $D_n * f$ et $F_n * f$?

16-2.7 Injectivité de la transformation de Fourier discrète

On appelle transformation de Fourier discrète l'application

$$\begin{aligned} \mathcal{C}_m^0([0,2\pi],\mathbb{C}) &\to \mathbb{C}^{\mathbb{Z}} \\ f &\mapsto \widehat{f} = (c_k(f))_{k\in\mathbb{Z}} \end{aligned}$$

On utilise ainsi la notation $\widehat{f}(k) = c_k(f)$. Cette application est linéaire, et n'est évidemment pas surjective, puisqu'elle prend ses valeurs dans l'ensemble des familles complexes indexées par $\mathbb{Z}$ de carré sommable. Une conséquence du théorème de Parseval est le

THÉORÈME 16-2.15 Si $f \in \mathcal{C}_m^0([0,2\pi],\mathbb{C})$ vérifie

$$\forall k \in \mathbb{Z} \qquad c_k(f) = 0$$

alors f est nulle sauf en un nombre fini de points. En particulier, la restriction de la transformation de Fourier à l'espace des fonctions continues sur $[0,2\pi]$ est injective.

> *Démonstration* : conséquence immédiate de l'égalité de Parseval-Bessel.

Une conséquence très importante de ce résultat est qu'une fonction continue est caractérisée par la suite de ses coefficients de Fourier : c'est la réponse (théorique) au problème de "synthèse harmonique" permettant souvent de caractériser la solution d'un problème par la suite de ses coefficients de Fourier, même lorsqu'on ne sait pas que la série de Fourier correspondante converge. Beaucoup de propriétés d'une fonction peuvent se lire sur ses coefficients de Fourier :

EXERCICE 16-2.16 Soit f une fonction continue $[0,2\pi] \to \mathbb{C}$. Montrer que f est restriction à $[0,2\pi]$ d'une fonction $\mathcal{C}^\infty$ 2π-périodique ssi ses coefficients de Fourier sont à décroissance rapide

$$\forall p \in \mathbb{N} \quad c_n(f) \underset{n\to\pm\infty}{=} o\left(\frac{1}{n^p}\right)$$

16-2.8 Cas des fonctions T-périodiques

Toute l'étude qui précède a été effectuée dans le cas de la 2π-périodicité. Pour une fonction f continue par morceaux et T-périodique, on aurait des résultats analogues, qu'on démontrerait aisément en utilisant la fonction

$$g(t) = f\left(\frac{T}{2\pi}t\right)$$

Les coefficients de Fourier de f sont, par définition, ceux de g. On vérifie par changement de variable dans les intégrales

$$c_k(f) = \frac{1}{T}\int_0^T f(t)\, e^{-\frac{2i\pi kt}{T}}\, dt$$

et de même

$$a_0(f) = \frac{1}{T}\int_0^T f(t)\,dt \quad \text{et, pour } k \geqslant 1$$
$$a_k(f) = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2k\pi}{T}t\right)dt,\ b_k(f) = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2k\pi}{T}t\right)dt$$

les fonctions exponentielles et trigonométriques intervenant dans ces calculs sont à présent T-périodiques. La série de Fourier de f est la série

$$a_0(f) + \sum_{k=1}^{+\infty}\left(a_k(f)\cos\left(\frac{2k\pi}{T}t\right) + b_k(f)\sin\left(\frac{2k\pi}{T}t\right)\right)$$

Les théorèmes qui précèdent (Dirichlet, Parseval etc...) se réécrivent aisément. En particulier, la formule de Parseval s'écrira

$$\frac{1}{T}\int_0^T |f(t)|^2 = |a_0(f)|^2 + \frac{1}{2}\sum_{n=1}^{+\infty}\left(|a_n(f)|^2 + |b_n(f)|^2\right)$$

EXERCICE 16-2.17 Reprendre l'exercice **16-2**.14 sur un segment $[a,b]$ quelconque.

Attention ! Lorsque les intégrations par parties sont possibles (*cf.* section **16-1.4**) on aura cette fois

$$c_k(f') = \frac{2ik\pi}{T}c_k(f)$$

résultat qu'on retiendra encore en dérivant **formellement** la série de Fourier de la fonction f.

16-2.9 Exemple d'utilisation

Les applications des séries de Fourier sont multiples : sommation de certaines séries, résolution de certaines équations différentielles ou aux dérivées partielles. C'est d'ailleurs pour étudier l'équation de la chaleur que Fourier introduisit les séries qui portent son nom. De nombreux phénomènes périodiques peuvent être modélisés par des séries de Fourier, celles-ci devenant ainsi un outil inévitable (en traitement du signal notamment). Nous traitons ici un exercice d'application (très élémentaire) à la recherche de solutions périodiques d'équations différentielles :

Soit $g : \mathbb{R} \to \mathbb{C}$, continue et T périodique. On cherche une condition nécessaire et suffisante pour qu'il existe au moins une solution T-périodique de l'équation différentielle

$$y'' + 4y = g$$

La théorie générale des équations différentielles linéaires montre que l'ensemble des solutions sur $\mathbb{R}$ de cette équation différentielle est un espace affine de dimension 2, inclus dans $C^2(\mathbb{R},\mathbb{C})$. Comme les solutions de l'équation homogène sont de la forme

$$t \mapsto \lambda\cos 2t + \mu\sin 2t$$

et sont toutes π-périodiques (plus petite période si la solution n'est pas nulle), il doit être clair que, si l'équation possède une solution T-périodique, celle-ci sera unique si $T \notin \pi\mathbb{Z}$. Dans le cas où une telle solution existe, et où $T \in \pi\mathbb{Z}$, toutes les solutions seront T-périodiques.

Analyse : si une telle solution f existe, comme f est $\mathcal{C}^2$, les coefficients de Fourier de $f'' + 4f$ sont donnés par

$$c_k(f'' + 4f) = \left(4 - \frac{4\pi^2k^2}{T^2}\right) c_k(f) = c_k(g)$$

- Si $T \notin \pi\mathbb{Z}$, le contenu de la parenthèse ne s'annule pas et on obtient

$$\forall k \in \mathbb{Z} \quad c_k(f) = \frac{c_k(g)}{\left(4 - \frac{4\pi^2k^2}{T^2}\right)}$$

  et l'on retrouve, par l'intermédiaire du théorème **16-2.15**, la remarque précédente : si f existe, elle est unique.

- Si $T \in \pi\mathbb{Z}$, il existe un entier $k_0 \in \mathbb{Z}$ avec $T = k_0\pi$. L'équation précédente montre que f ne peut exister que si

$$c_{k_0}(g) = c_{-k_0}(g) = 0$$

  les coefficients $c_{k_0}(f)$ et $c_{-k_0}(f)$ sont alors "indéterminés", les autres étant encore donnés par

$$\forall k \neq \pm k_0 \quad c_k(f) = \frac{c_k(g)}{\left(4 - \frac{4\pi^2k^2}{T^2}\right)}$$

Synthèse : supposons les conditions nécessaires précédentes vérifiées, et recherchons à présent une fonction f possédant les coefficients de Fourier données par les formules précédentes (dans le cas $T = k_0\pi$, on prendra par exemple $c_{k_0}(f)$ et $c_{-k_0}(f)$ nuls). Une telle fonction existe : comme les coefficients de Fourier de g tendent vers 0 à l'infini, on a

$$c_k(f) \underset{k\to\pm\infty}{=} o\left(\frac{1}{k^2}\right)$$

ce qui prouve la sommabilité de la famille de ces coefficients, et permet de définir

$$f(t) = \sum_{k\in\mathbb{Z}} \frac{c_k(g)}{\left(4 - \frac{4\pi^2k^2}{T^2}\right)} e^{\frac{2ik\pi}{T}t}$$

(avec ici la convention $\frac{0}{0} = 0$). En travaillant avec les sinus et cosinus, on obtient une série normalement convergente, à laquelle on peut appliquer le théorème **16-1.17** : f est une fonction continue de période T, dont les coefficients de Fourier répondent à la question, vérifiant

$$\forall k \quad c_k(f) = \frac{c_k(g)}{\left(4 - \frac{4\pi^2k^2}{T^2}\right)}$$

Pour terminer tout à fait la synthèse, il faut prouver que f est C^2, puisqu'alors l'égalité

$$\left(4-\frac{4\pi^2k^2}{T^2}\right)c_k(f)=c_k(g)$$

donnera

$$\forall k \quad c_k(f''+4f)=c_k(g)$$

et $f''+4f=g$ par injectivité de la transformée de Fourier. Le caractère C^1 de f ne pose pas de problème, car on peut facilement dériver terme à terme la relation

$$f(t)=\sum_{k\in\mathbb{Z}}\frac{c_k(g)}{\left(4-\frac{4\pi^2k^2}{T^2}\right)}e^{\frac{2ik\pi}{T}t}$$

par convergence normale de la série des dérivées puisque

$$\frac{c_k(g)}{\left(4-\frac{4\pi^2k^2}{T^2}\right)}\frac{2ik\pi}{T}=O\left(\frac{c_k(g)}{k}\right)=O\left(|c_k(g)|^2+\frac{1}{k^2}\right)$$

On pourrait de même dériver deux fois sans problème si on supposait g continue et C^1 par morceaux. Dans le cas où l'on suppose simplement g continue, on pourrait utiliser par exemple le théorème de Fejér : les sommes de Fejér de la fonction f vérifient, par linéarité de l'équation

$$(\Sigma_n^f)''=-4\Sigma_n^f+\Sigma_n^g$$

La suite de ces dérivées secondes converge donc uniformément sur $\mathbb{R}$ vers $-4f+g$. Comme la suite des dérivées premières converge simplement vers f' (pourquoi ?) et que $\left(\Sigma_n^f\right)_{n\in\mathbb{N}}$ converge simplement vers f, on conclut facilement au caractère C^2 de f.

16-3 Exercices

EXERCICE 16-3.1 Soit $a\in\mathbb{C}-[-1,1]$. Calculer $\displaystyle\int_0^\pi\frac{\cos nt}{a-\cos t}dt$.

EXERCICE 16-3.2 α étant fixé dans $\mathbb{R}^{*+}$, on considère l'application ψ_α définie par

$$\psi_\alpha=\sum_{k\in\mathbb{Z}}e^{-\frac{\alpha}{2}(x-2k\pi)^2}$$

Existence de ψ_α. Montrer qu'elle est C^1 et 2π-périodique. Calculer ses coefficients de Fourier. Pour α et β fixés dans $\mathbb{R}$, on définit φ par

$$\varphi(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\psi_\alpha(x-t)\psi_\beta(t)dt$$

Montrer qu'il existe $b\in\mathbb{R}$ et $\gamma\in\mathbb{R}^{*+}$ avec $\varphi=b\psi_\gamma$ (on pourra calculer les coefficients de φ).

EXERCICE 16-3.3 Soient $\alpha, \beta, \gamma \in \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}$ 2π-périodique et paire définie sur $[0, \pi]$ par

$$f(x) = \alpha x^2 + \beta x + \gamma$$

Que peut-on dire de la série de Fourier de f ? En déduire $\sum \frac{1}{n^2}$.

EXERCICE 16-3.4 On désigne par S la somme de la série de fonctions

$$u_n(x) = (-1)^n \frac{\cos nx}{n^2(a^2 + n^2)}$$

Montrer que S est de classe C^2 et vérifie sur $[-\pi, \pi]$

$$S''(x) - a^2 S(x) = \frac{\pi^2}{12} - \frac{x^2}{4}$$

En déduire la valeur de S.

EXERCICE 16-3.5 Soit f une application 2π-périodique et α un réel > 0 tels que

$$\forall x, y \in \mathbb{R} \quad |f(x) - f(y)| \leqslant |x - y|^\alpha$$

Calculer $\int_0^{2\pi} |f(x+h) - f(x-h)|^2 dx$. En déduire que

$$\sum_{n=-\infty}^{+\infty} |c_n|^2 \sin^2 nh \ \leqslant\ A|h|^{2\alpha}$$

où A est une constante à déterminer. En déduire une majoration de $\sum_{N+1}^{2N} |c_n|^2$. Si $\alpha > \frac{1}{2}$ que peut-on dire de la convergence de la série de Fourier de f ?

EXERCICE 16-3.6 Calculer les sommes des séries $\sum_{n=1}^{+\infty} \frac{\sin n\theta}{n}$ et $\sum_{n=1}^{+\infty} \frac{\cos n\theta}{n}$ lorsqu'elles existent.

EXERCICE 16-3.7 Soit E l'ensemble des fonctions C^∞ et 2π périodiques de $\mathbb{R}$ dans $\mathbb{C}$.

1. Soit $(a_n) \in \mathbb{C}^{\mathbb{Z}}$. Donner une condition nécessaire et suffisante pour que cette suite soit la suite des coefficients de Fourier d'une fonction de E.
2. Soit $\theta \in 2\pi\mathbb{Q}$ et L_θ l'application de E dans lui-même définie par

$$\forall f \in E \quad L_\theta f(x) = f(x + \theta) - f(x)$$

Caractériser les éléments de l'image de L_θ.

EXERCICE 16-3.8 Soit $(b_n)_{n \in \mathbb{N}^*}$ une suite strictement positive, décroissante et tendant vers 0.

1. Quel est le domaine de convergence de la série $\sum_{n=1}^{+\infty} b_n \sin nx$?

2. On suppose $nb_n \to 0$ et on pose $\varepsilon_n = \sup_{k \geqslant n}(kb_k)$. Si $x \in]0,\pi]$ et p désigne la partie entière de $\frac{\pi}{x}$, en écrivant $R_n(x) = \sum_{n}^{+\infty} b_k \sin kx = \sum_{n}^{n+p-1} b_k \sin kx + \sum_{n+p}^{+\infty} b_k \sin kx$, montrer que $|R_n| \leqslant (\pi+2)\varepsilon_n$. Conclure.

3. Montrer que, si la série converge uniformément sur $[0,\pi]$, alors $nb_n \to 0$.

EXERCICE 16-3.9 Soit $\lambda \in \mathbb{R}$ et u une application π-périodique $\mathbb{R} \to \mathbb{R}$ vérifiant l'équation

$$(E) \qquad u'' + (\lambda - \cos 2x)u = 0$$

1. Trouver une relation de récurrence entre les coefficients de Fourier $\left(c_n(u)\right)_{n\in\mathbb{Z}}$.

2. On note E_1 l'espace vectoriel des suites complexes indexées par $\mathbb{Z}$, bornées et vérifiant la relation de récurrence du (1). Montrer que la dimension de E_1 est inférieure à 1, et que les suites de E_1 sont soit paires, soit impaires.

3. On suppose $\dim(E_1) = 1$. Montrer que (E) possède une solution π-périodique non triviale.

EXERCICE 16-3.10 Soit f 2π-périodique C^1 de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\int_0^{2\pi} f(x)dx = 0$. Montrer l'inégalité $\int_0^{2\pi} (f(t))^2\, dt \leqslant \int_0^{2\pi} \left(f'(t)\right)^2 dt$, et étudier les cas d'égalité.

EXERCICE 16-3.11 Donner le développement en série de Fourier de $f(t) = |\sin t|$.

EXERCICE 16-3.12 On appelle E l'ensemble des fonctions continues et 2π-périodiques de $\mathbb{R}$ dans $\mathbb{R}$, et Ω l'application définie sur E par :

$$\forall g \in E,\ \forall x \in \mathbb{R},\ \Omega(g)(x) = \int_{-\pi}^{+\pi} f(x-t)\, g(t)dt$$

Montrer que Ω est un endomorphisme de E et déterminer son spectre et ses sous-espaces propres.

EXERCICE 16-3.13 Soit k un réel strictement positif. Trouver toutes les fonctions f 2π-périodiques et C^∞ sur $\mathbb{R}$ telles qu'il existe $M > 0$ tel que

$$\forall n \in \mathbb{N} \quad \forall x \in \mathbb{R} \quad |f^{(n)}(x)| \leqslant Mk^n$$

EXERCICE 16-3.14 Calculer la somme $f(x)$ de la série de terme général $\frac{\sin nx}{n}$. On en déduit une fonction g impaire et 2π-périodique, continue de telle sorte que g est affine sur $[0,1]$ et que f et g coïncident sur $[1,\pi]$. En utilisant g, montrer que

$$\sum_{n=1}^{+\infty} \frac{\sin^2 n}{n^2} = \sum_{n=1}^{+\infty} \frac{\sin n}{n}$$

Calculer ensuite $\sum_{n=1}^{+\infty} \frac{\sin^2 n}{n^4}$.

EXERCICE 16-3.15 Formule sommatoire de Poisson
On note $S(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}, \mathbb{C}) \mid \forall k, l \;\; x^l f^{(k)}(x) = o(1) \;\; (x \to \pm\infty)\}$ l'espace des fonctions C^∞ à décroissance rapide. Montrer que, pour $f \in S(\mathbb{R})$ la série $\sum_{n \in \mathbb{Z}} f(x+n)$ converge uniformément sur tout compact de $\mathbb{R}$. Montrer que la fonction ainsi définie est C^∞ et calculer ses coefficients de Fourier à l'aide de la transformée de Fourier $\widehat{f}$ de f

$$\widehat{f}(x) = \int_{-\infty}^{+\infty} f(t) e^{-2i\pi tx}\, dt$$

En déduire que $\sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \widehat{f}(n)$.

EXERCICE 16-3.16 Soit f une fonction continue par morceaux nulle hors d'un segment $[-A, A]$ et $\widehat{f}$ sa transformée de Fourier

$$\widehat{f}(x) = \int_{-A}^{A} f(t) e^{-2i\pi tx}\, dt$$

On suppose qu'il existe un réel $B > 0$ tel que $\widehat{f}$ soit nulle hors de $[-B, B]$. Montrer que f est nulle sauf en un nombre fini de points. (Indication : prolonger f par $2A$-périodicité).

EXERCICE 16-3.17 Soit f une application continue de $\mathbb{R}$ dans $\mathbb{R}$ de période 2π. Soit a_n et b_n ses coefficients de Fourier. Pour $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$, on pose :

$$D_n(x) = \sum_{k=1}^{n} \sin kx \quad \text{et} \quad S_n(x) = \sum_{k=1}^{n} (a_k \sin kx \; - \; b_k \cos kx)$$

Soit $\psi_x(t) = f(x-t) - f(x+t)$. Donner une expression intégrale de $S_n(x)$ faisant intervenir ψ_x et D_n. Que peut-on dire de la suite $S_n(x)$ lorsque la fonction $t \mapsto \dfrac{\psi_x(t)}{t}$ est intégrable sur $]0, \pi]$?

Chapitre 17

Groupes. Anneau $\mathbb{Z}/n\mathbb{Z}$

Dans un chapitre précédent, nous avons vu comment la notion d'anneau principal permettait d'unifier les théories du P.G.C.D et du P.P.C.M sur $\mathbb{Z}$ et $\mathbb{K}[X]$, si $\mathbb{K}$ est un corps commutatif. Dans ce chapitre, après avoir précisé quelques notions sur les groupes, nous utiliserons certains de ces résultats pour obtenir des théorèmes sur la divisiblité dans $\mathbb{Z}$. Les groupes utilisés seront essentiellement le groupe additif et le groupe des unités de l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ des classes de congruence des entiers modulo n.

17-1 Structure de groupe

17-1.1 Groupe, morphisme de groupe

17-1.1.1 Structure de groupe

Rappelons qu'un groupe est un ensemble non vide $\mathbb{G}$ muni d'une loi de composition interne, que nous noterons en général multiplicativement

$$\mathbb{G}^2 \ni (x, y) \mapsto x.y \in \mathbb{G} \text{ (noté en général } xy)$$

tel que

- La loi de composition interne est associative.
- Elle possède un élément neutre, que nous noterons en général $1_{\mathbb{G}}$.

- Tout élément de $\mathbb{G}$ est inversible pour la loi de composition interne :

$$\forall x \in \mathbb{G} \quad \exists x' \in \mathbb{G} \qquad xx' = x'x = 1_{\mathbb{G}}$$

L'élément x' vérifiant cette égalité[1] pour $x \in \mathbb{G}$ est appelé inverse de x. Nous le noterons en général x^{-1}. Lorsque la loi de composition est de plus commutative, on dit que le groupe $(\mathbb{G}, .)$ est commutatif ou abélien.

Historiquement, la notion de groupe a été introduite à partir de problèmes géométriques, en considérant des ensembles de transformations d'un ensemble X stables par composition et possédant pour cette loi, les propriétés précédentes (isométries du plan, homothéties-translations etc...). Il est remarquable que de très nombreux résultats sur la structure d'un groupe "abstrait" s'obtiennent par des considérations "géométriques", en faisant "opérer le groupe sur des ensembles" (*cf.* section **17-2**). Contentons-nous pour le moment de donner l'exemple de référence en théorie des groupes :

EXEMPLE 17-1.1 Si X est un ensemble non vide, l'ensemble Σ_X des bijections de X sur lui-même est un groupe pour la composition des applications. Le groupe $(\Sigma_X, \circ)$ est appelé groupe des permutations de l'ensemble X. Si l'ensemble X est fini de cardinal n, on sait que Σ_X est fini, de cardinal $n!$.

17-1.1.2 Règles de calcul dans un groupe

On a évidemment, pour $x, y \in \mathbb{G}$

$$(xy)^{-1} = y^{-1}x^{-1}$$

Pour $x \in \mathbb{G}$ et $n \in \mathbb{Z}$, nous notons

$$x^n = \begin{cases} \underbrace{x.x.\ldots.x}_{n \text{ fois}} & \text{si} \quad n \in \mathbb{N}^* \\ \underbrace{x^{-1}\ldots x^{-1}}_{-n \text{ fois}} & \text{si} \quad n < 0 \\ 1_{\mathbb{G}} & \text{si} \quad n = 0 \end{cases}$$

On vérifie alors aisément que

$$\forall n, p \in \mathbb{Z} \quad x^n.x^p = x^{n+p}$$

Il faut par contre garder à l'esprit que, pour $x, y \in \mathbb{G}$ et $n \in \mathbb{Z}$, l'égalité

$$(xy)^n = x^n y^n$$

n'est en général pas vérifiée ! Elle l'est cependant si x et y commutent.

[1] Il est en effet unique. Si x' et x'' vérifient la relation, puisque la loi de composition est associative, nous avons

$$x' = x'.(xx'') = (x'x).x'' = x''$$

REMARQUE 17-1.2 Lorsque la loi de groupe est notée additivement, il est en général implicite que le groupe est commutatif[2]. On note alors $0_{\mathbb{G}}$ l'élément neutre et $-x$ le symétrique de $x \in \mathbb{G}$ (on préfèrera dans ce cas la dénomination "opposé de x" à "inverse" de x) et la notation exponentielle est alors remplacée par

$$\forall x \in \mathbb{G} \quad \forall n \in \mathbb{G} \quad n.x = \begin{cases} \underbrace{x + x + \ldots + x}_{n \text{ fois}} & \text{si} \quad n \in \mathbb{N}^* \\ \underbrace{(-x) + \ldots + (-x)}_{-n \text{ fois}} & \text{si} \quad n < 0 \\ 0_{\mathbb{G}} & \text{si} \quad n = 0 \end{cases}$$

Passer d'un type de notation à l'autre demande un peu d'attention. Il faudra savoir traduire en notation additive les définitions qui seront exposées avec la notation multiplicative (en particulier, tout ce qui est relatif à la notion d'ordre d'un élément dans un groupe).

PROPOSITION 17-1.3 Si $(\mathbb{G}, .)$ est un groupe et $a \in \mathbb{G}$, les applications de $\mathbb{G}$ dans lui-même

$$x \mapsto a\,x \quad \textbf{et} \quad x \mapsto x\,a$$

(respectivement appelées translation à gauche et à droite par a) sont des bijections de $\mathbb{G}$ dans lui-même.

Démonstration : Il suffit de voir que, pour $b \in \mathbb{G}$, l'équation d'inconnue $x \in \mathbb{G}$

$$a\,x = b$$

possède une solution unique

$$x = a^{-1}b$$

ce qui découle facilement de l'associativité de la l.c.i et de la définition de a^{-1}. De même

$$x\,a = b \Leftrightarrow x = b\,a^{-1} \quad \blacksquare$$

[2]Si $(\mathbb{G}, +)$ est un tel groupe, on peut munir $\mathbb{G}$ d'une multiplication externe

$$\mathbb{Z} \times \mathbb{G} \to \mathbb{G} \qquad (n, x) \mapsto n.x$$

$(\mathbb{G}, +, .)$ possède alors toutes les propriétés d'un espace vectoriel, à ceci près que $(\mathbb{Z}, +, \times)$ est un anneau intègre, mais n'est pas un corps. On dit alors que $(\mathbb{G}, +, .)$ est un $\mathbb{Z}$-module. On y calcule un peu comme dans un espace vectoriel, mais les manipulations de combinaisons linéaires sont plus délicates. En particulier, pour $x, y \in \mathbb{G}$, l'existence d'une "relation de liaison"

$$n.x + m.y = 0_{\mathbb{G}}$$

avec $n, m \in \mathbb{Z}$ et $n \neq 0$ ne permet pas en général d'obtenir x comme "combinaison linéaire" de y : dans le groupe $(\mathbb{Z}, +)$, prendre $x = 2$ et $y = -3$, avec par exemple $n = 3$ et $m = 2$. La propriété

$$n.x = 0_{\mathbb{G}} \Rightarrow n = 0_{\mathbb{Z}} \text{ ou } x = 0_{\mathbb{G}}$$

n'est également plus vérifiée en général. Elle n'est par exemple certainement pas valable dans un groupe $(\mathbb{G}, +)$ fini non réduit à l'élément neutre (exercice).

En particulier, tout élément d'un groupe est régulier, c'est-à-dire simplifiable à droite et à gauche

$$\forall a, x, y \in \mathbb{G} \qquad a\,x = a\,y \Leftrightarrow x = y \Leftrightarrow x\,a = y\,a$$

COROLLAIRE 17-1.4 Si $\mathbb{G}$ est un groupe et $x, y \in \mathbb{G}$, l'égalité $xy = 1_{\mathbb{G}}$ entraîne que y est l'inverse de x (sans qu'il soit nécessaire de vérifier $yx = 1_{\mathbb{G}}$).

EXERCICE 17-1.5 Soit $\mathbb{X}$ un ensemble **fini** muni d'une loi de composition interne associative pour laquelle tout élément de $\mathbb{X}$ est régulier. Montrer que $\mathbb{X}$ muni de cette l.c.i est un groupe.

EXERCICE 17-1.6 Montrer que dans un groupe $(\mathbb{G}, .)$, l'équation $x^2 = x$ caractérise l'élément neutre.

17-1.1.3 Produit de groupes

Soient $(\mathbb{G}_1, .)$ et $(\mathbb{G}_2, .)$ deux groupes (où, comme d'habitude, nous notons les lois multiplicativement, mais ces deux lois sont en général différentes. Nous les noterons différemment uniquement lorsqu'il risque d'y avoir des confusions). Le produit cartésien $\mathbb{G}_1 \times \mathbb{G}_2$ peut être muni d'une structure "naturelle" de groupe, appelée structure de groupe produit de $(\mathbb{G}_1, .)$ et $(\mathbb{G}_2, .)$. Cette loi (que nous notons toujours multiplicativement) est simplement définie par

$$\forall x_1, y_1 \in \mathbb{G}_1 \quad \forall x_2, y_2 \in \mathbb{G}_2 \qquad (x_1, x_2) \,.\, (y_1, y_2) = (x_1 y_1, x_2 y_2)$$

On vérifie aisément que $(1_{\mathbb{G}_1}, 1_{\mathbb{G}_2})$ est élément neutre de $\mathbb{G}_1 \times \mathbb{G}_2$ pour cette loi produit et que

$$\forall x_1 \in \mathbb{G}_1 \quad \forall x_2 \in \mathbb{G}_2 \qquad (x_1, x_2)^{-1} = \left(x_1^{-1}, x_2^{-1}\right)$$

On généralise évidemment au produit d'un nombre fini de groupes.

17-1.1.4 Morphisme de groupes

Rappelons la définition :

DÉFINITION 17-1.7 *Si $(\mathbb{G}_1, .)$ et $(\mathbb{G}_2, .)$ sont deux groupes. Une application $\varphi : \mathbb{G}_1 \to \mathbb{G}_2$ est un morphisme de groupes (sous-entendu : pour les lois de groupes!) si et seulement si*

$$\forall x, y \in \mathbb{G}_1 \qquad \varphi(xy) = \varphi(x)\,\varphi(y)$$

Si, de plus, φ est bijective, on dit que φ est un isomorphisme. Un automorphisme d'un groupe $(\mathbb{G}, .)$ est un isomorphisme de $\mathbb{G}$ sur lui-même.

Les propriétés suivantes sont conséquence immédiate de la définition :

PROPOSITION 17-1.8 Si $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$ est un morphisme de groupe, on a

$$\varphi(1_{\mathbb{G}_1}) = 1_{\mathbb{G}_2} \quad \textbf{et} \quad \forall x \in \mathbb{G}_1 \quad \varphi\left(x^{-1}\right) = [\varphi(x)]^{-1}$$

Démonstration : Il suffit de transporter par φ les égalités

$$1_{\mathbb{G}_1}^2 = 1_{\mathbb{G}_1} \quad \text{et} \quad xx^{-1} = 1_{\mathbb{G}_1} \quad \blacksquare$$

PROPOSITION 17-1.9 Si $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$ et $\psi : (\mathbb{G}_2, .) \to (\mathbb{G}_3, .)$ sont des morphismes de groupes, leur composé $\psi \circ \varphi$ est un morphisme de $(\mathbb{G}_1, .) \to (\mathbb{G}_3, .)$. En particulier, le composé de deux isomorphismes de groupes est un isomorphisme.

PROPOSITION 17-1.10 Si $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$ est un isomorphisme de groupes, la bijection réciproque $\varphi^{-1} : (\mathbb{G}_2, .) \to (\mathbb{G}_1, .)$ est un isomorphisme.

COROLLAIRE 17-1.11 Si $(\mathbb{G}, .)$ est un groupe, l'ensemble $\mathrm{Aut}(\mathbb{G})$ des automorphismes de $\mathbb{G}$ est non vide (il contient au moins $id_{\mathbb{G}}$). La composition des applications est une loi interne dans $\mathrm{Aut}(\mathbb{G})$, et

$$(\mathrm{Aut}(\mathbb{G}), \circ) \text{ est un groupe}$$

EXERCICE 17-1.12 Déterminer tous les morphismes de $(\mathbb{Z}, +)$ dans lui-même, de $(\mathbb{Z}, +)$ dans $(\mathbb{Q}, +)$, de $(\mathbb{Q}, +)$ dans $(\mathbb{Z}, +)$.

On dit qu'un groupe $(\mathbb{G}_1, .)$ est **isomorphe** à un groupe $(\mathbb{G}_2, .)$ s'il existe (au moins) un isomorphisme $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$. Les propositions précédentes montrent que l'on définit ainsi une relation réflexive, symétrique et transitive sur la classe des groupes. Il faut bien comprendre la signification de l'affirmation "$(\mathbb{G}_1, .)$ et $(\mathbb{G}_2, .)$ sont isomorphes" : cela signifie qu'il existe un procédé (en l'occurence un isomorphisme $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$) qui permet de transporter les propriétés de la loi de composition de $\mathbb{G}_1$ sur des propriétés analogues de la loi de $\mathbb{G}_2$. Si on décide d'identifier un élément x de $\mathbb{G}_1$ et son image $\varphi(x) \in \mathbb{G}_2$, on calcule dans $\mathbb{G}_2$ comme dans $\mathbb{G}_1$. Par exemple, une propriété de $\mathbb{G}_1$ qui s'écrirait

$$\exists a \in \mathbb{G}_1 \quad \forall x \in \mathbb{G}_1 \quad \exists n \in \mathbb{Z} \quad x = a^n$$

sera vérifiée par tout groupe $\mathbb{G}_2$ isomorphe à $\mathbb{G}_1$. Si $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$ est un isomorphisme, tout élément de $\mathbb{G}_2$ pourra s'écrire comme puissance de $b = \varphi(a)$.

EXEMPLE 17-1.13 Si X et Y sont deux ensembles équipotents (c'est-à-dire qu'il existe une bijection $f : X \to Y$), les groupes de permutation $(\Sigma_X, \circ)$ et $(\Sigma_Y, \circ)$ sont isomorphes. Expliciter un isomorphisme entre ces deux groupes.

17-1.2 Sous-groupe

17-1.2.1 Définition

Si $(\mathbb{G}, .)$ est un groupe, une partie non vide $\mathbb{H} \subset \mathbb{G}$ est un sous-groupe de $\mathbb{G}$ si et seulement si :

1. $\mathbb{H}$ est stable pour la loi du groupe $\mathbb{G}$:

$$\forall x, y \in \mathbb{H} \quad xy \in \mathbb{H}$$

2. $\mathbb{H}$ muni de la restriction (à $\mathbb{H}\times\mathbb{H}$) de la loi interne est lui-même un groupe :

$$(\mathbb{H},.) \text{ est un groupe}$$

L'élément neutre[3] de $(\mathbb{H},.)$ est évidemment $1_\mathbb{G}$ (seul élément de $\mathbb{G}$, et *a fortiori* de $\mathbb{H}$ vérifiant $x^2 = x$). De même, si $x \in \mathbb{H}$, son inverse dans le groupe $(\mathbb{H},.)$ est égal à son inverse dans $\mathbb{G}$ (seul élément y de $\mathbb{G}$ vérifiant $xy = 1_\mathbb{G}$). On en déduit aisément les caractérisations des sous-groupes :

PROPOSITION 17-1.14 $\mathbb{H} \subset \mathbb{G}$ **est un sous-groupe de** $(\mathbb{G},.)$ **si et seulement si** $\mathbb{H} \neq \emptyset$, $\mathbb{H}$ **est stable pour la loi du groupe**

$$\forall x,y \in \mathbb{H} \quad xy \in \mathbb{H}$$

et $\mathbb{H}$ **est stable pour l'opération de passage à l'inverse**

$$\forall x \in \mathbb{H} \qquad x^{-1} \in \mathbb{H}$$

Démonstration : Ces deux conditions sont nécessaires pour que $\mathbb{H}$ soit un sous-groupe, comme on vient de le voir. Elles sont évidemment suffisantes : lorsqu'elles sont vérifiées, $\mathbb{H}$ contient $1_\mathbb{G}$ (puisque $\mathbb{H}$ est non vide, si $a \in \mathbb{H}$, on a $a^{-1} \in \mathbb{H}$ et $aa^{-1} \in \mathbb{H}$ d'après les propriétés de stabilité). Il est alors clair que $(\mathbb{H},.)$ est un groupe. ■

On peut parfois préférer la version plus condensée (exercice) :

PROPOSITION 17-1.15 $\mathbb{H} \subset \mathbb{G}$ **est un sous-groupe de** $(\mathbb{G},.)$ **si et seulement si** $\mathbb{H} \neq \emptyset$ **et**

$$\forall x,y \in \mathbb{H} \quad xy^{-1} \in \mathbb{H}$$

REMARQUE 17-1.16 La notion de sous-groupe est très souvent utilisée pour montrer qu'un ensemble $\mathbb{X}$ muni d'une loi de composition interne $*$ est un groupe : il suffit souvent de le "plonger" dans un groupe "bien connu" (avec une loi qui est évidemment prolongement de $*$) et de vérifier que $(\mathbb{X}, *)$ est un sous-groupe de ce groupe de référence[4]. On évite ainsi des vérifications fastidieuses, notamment sur l'associativité de la loi.

EXERCICE 17-1.17 Montrer que l'ensemble des matrices de $\mathcal{M}_4(\mathbb{R})$ écrites en blocs sous la forme

$$\begin{pmatrix} \multicolumn{2}{c}{\multirow{2}{*}{A}} & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

avec $A \in \mathcal{GL}_2(\mathbb{R})$ est un groupe pour la multiplication des matrices. Quel est son élément neutre ?

EXERCICE 17-1.18 Si $\mathbb{H}_1$ et $\mathbb{H}_2$ sont deux sous-groupes de $(\mathbb{G},.)$, montrer que

$$\mathbb{H}_1 \cup \mathbb{H}_2 \text{ sous-groupe} \Leftrightarrow \mathbb{H}_1 \subset \mathbb{H}_2 \text{ ou } \mathbb{H}_2 \subset \mathbb{H}_1$$

[3]Ainsi, pour vérifier qu'une partie d'un groupe est un sous-groupe, on s'assurera le plus souvent que cette partie est non vide en montrant qu'elle contient l'élément neutre.

[4]Attention : ce groupe de référence doit être un ... groupe ! Il n'y a rien de plus désastreux que d'essayer de prouver, par exemple, qu'un ensemble de matrices carrées est un groupe pour la multiplication en faisant référence à une structure de groupe sur $(\mathcal{M}_n(\mathbb{K}), \times)$!

17-1.2.2 Exemples

Les sous-groupes du groupe $(\mathbb{Z},+)$ sont identiques aux idéaux de l'anneau $(\mathbb{Z},+,\times)$. Cela est dû au fait que, dans $\mathbb{Z}$, la multiplication est directement liée à l'addition : pour $n \in \mathbb{N}^*$ et $m \in \mathbb{Z}$, le produit nm n'est autre que $m+\cdots+m$ (n fois). Le théorème **5-1**.19, conséquence de l'existence de la division euclidienne dans $\mathbb{Z}$, s'écrit aussi :

PROPOSITION 17-1.19 Si $\mathbb{H}$ est un sous-groupe de $(\mathbb{Z},+)$, il existe un unique $m \in \mathbb{N}$ avec $\mathbb{H} = m\mathbb{Z}$.

Un morphisme de groupes $\varphi : (\mathbb{G}_1,.) \to (\mathbb{G}_2,.)$ permet de "transporter les sous-groupes" par images directes et réciproques. Plus précisément :

PROPOSITION 17-1.20 Si $\varphi : (\mathbb{G}_1,.) \to (\mathbb{G}_2,.)$ est un morphisme de groupes,

$$\mathbb{H}_1 \text{ sous-groupe de } \mathbb{G}_1 \Rightarrow \varphi(\mathbb{H}_1) \text{ sous-groupe de } \mathbb{G}_2$$

$$\mathbb{H}_2 \text{ sous-groupe de } \mathbb{G}_2 \Rightarrow \varphi^{-1}(\mathbb{H}_2) \text{ sous-groupe de } \mathbb{G}_1$$

En particulier, l'image de φ et le noyau de φ

$$\operatorname{Im}\varphi = \varphi(\mathbb{G}_1) \quad \textbf{et} \quad \ker\varphi = \varphi^{-1}(1_{\mathbb{G}_2})$$

sont des sous-groupes respectifs de $\mathbb{G}_2$ et $\mathbb{G}_1$.

Démonstration : Exercice. On remarquera évidemment que, comme on a toujours $\varphi(1_{\mathbb{G}_1}) = 1_{\mathbb{G}_2}$, l'image réciproque d'un sous-groupe de $\mathbb{G}_2$ par le morphisme φ contient $1_{\mathbb{G}_1}$, et n'est donc pas vide. Plus précisément, cette image réciproque contient le noyau de φ. ■

EXERCICE 17-1.21 Montrer que la correspondance

$$\mathbb{H}_1 \mapsto \varphi(\mathbb{H}_1)$$

réalise une bijection entre l'ensemble des sous-groupes de $\mathbb{G}_1$ contenant $\ker\varphi$ et l'ensemble des sous-groupes de $\operatorname{Im}\varphi$.

Le noyau d'un morphisme de groupes permet de caractériser le défaut d'injectivité de ce morphisme. Plus le noyau est "gros", plus il y a dans le groupe de départ d'éléments ayant la même image. Plus précisément, si $\varphi : (\mathbb{G}_1,.) \to (\mathbb{G}_2,.)$ est un morphisme de groupes, si x et $y \in \mathbb{G}_1$, on a

$$\varphi(x) = \varphi(y) \Leftrightarrow [\varphi(x)]^{-1}\varphi(y) = \varphi(x^{-1}y) = 1_{\mathbb{G}_2} \Leftrightarrow x^{-1}y \in \ker\varphi \Leftrightarrow y \in x.\ker\varphi$$

en notant[5]

$$x.\ker\varphi = \{x.z,\ z \in \ker\varphi\}$$

le sous-ensemble de $\mathbb{G}_1$ obtenu en faisant opérer la translation à gauche associée à x sur le noyau de φ. On appelle classe de x à gauche modulo $\ker\varphi$ cet ensemble qui contient donc tous les antécédents de $\varphi(x)$ par φ.

[5]Avec une loi additive, ce serait évidemment

$$x + \ker\varphi$$

notation que nous avons déjà rencontrée en algèbre linéaire : un morphisme d'espaces vectoriels est en particulier un morphisme de groupes additifs.

EXERCICE 17-1.22 Avec les notations qui précèdent, montrer que

$$\forall x \in \mathbb{G}_1 \quad x.\ker\varphi = \ker\varphi.x$$

ce qui peut aussi s'écrire $x^{-1}.\ker\varphi.x = \ker\varphi$. Les classes à gauche et à droite de x modulo $\ker\varphi$ sont donc égales. Ce résultat n'a évidemment d'intérêt que lorsque le groupe $\mathbb{G}_1$ n'est pas commutatif.

PROPOSITION 17-1.23 Un morphisme $\varphi : (\mathbb{G}_1, .) \to (\mathbb{G}_2, .)$ est injectif si et seulement si

$$\ker\varphi = \{1_{\mathbb{G}_1}\}$$

17-1.2.3 Sous groupe, groupe engendré par une partie

Nous suivons la même démarche que pour les sous-espaces vectoriels, les idéaux d'un anneau commutatif etc... Il s'agit ici de construire le plus petit sous-groupe (pour la relation d'inclusion) d'un groupe $(\mathbb{G}, .)$ contenant une partie $F \subset \mathbb{G}$. On l'obtient grâce au théorème, dont la démonstration est laissée en exercice :

THÉORÈME 17-1.24 Si $(\mathbb{H}_i)_{i\in I}$ est une famille non vide de sous-groupes de $(\mathbb{G}, .)$, l'intersection $\bigcap_{i\in I} \mathbb{H}_i$ est un sous-groupe de $\mathbb{G}$.

Si F est une partie quelconque de $\mathbb{G}$, l'ensemble

$$\mathcal{F} = \{\mathbb{H} \mid \mathbb{H} \text{ sous-groupe de } \mathbb{G} \text{ vérifiant } F \subset \mathbb{H}\}$$

est non vide, puisqu'il contient au moins $\mathbb{G}$. L'intersection de tous les éléments de $\mathcal{F}$ est donc un sous-groupe de $\mathbb{G}$. C'est par construction le plus petit sous-groupe de $\mathbb{G}$ contenant F. On l'appelle **sous-groupe de $\mathbb{G}$ engendré par F**. On le note

$$\mathrm{Gr}(F) = \bigcap_{\mathbb{H}\in\mathcal{F}} \mathbb{H}$$

On rencontre également la notation $\mathrm{Gr}(F) = \langle F\rangle$.

Lorsque nous avons rencontré la notion analogue de sous-espace vectoriel engendré par une famille de vecteurs d'un e.v. nous avons caractérisé les éléments de ce s.e.v. comme étant les combinaisons linéaires des éléments de la famille, les règles de calcul sur les combinaisons linéaires étant assez simples. Dans le cas des groupes, on a un résultat du même type. Mais les calculs sont moins aisés, en particulier en cas de non commutativité :

THÉORÈME 17-1.25 Dans un groupe $(\mathbb{G}, .)$, on a évidemment $\mathrm{Gr}(\emptyset) = \{1_{\mathbb{G}}\}$. Si $F \subset \mathbb{G}$ est non vide, en notant

$$F' = F \cup F^{-1} = F \cup \{x^{-1} \mid x \in F\}$$

l'ensemble obtenu en "symétrisant" F, le sous-groupe engendré par F est exactement l'ensemble des produits obtenus à partir des éléments de F' :

$$\mathrm{Gr}(F) = \{x \in \mathbb{G} \mid \exists n \in \mathbb{N}^* \quad \exists (x_i)_{1\leqslant i\leqslant n} \in F'^n \quad x = x_1x_2\cdots x_n\}$$

Démonstration : Notons

$$\mathbb{F} = \left\{x \in \mathbb{G} \mid \exists n \in \mathbb{N}^* \quad \exists (x_i)_{1 \leqslant i \leqslant n} \in F^n \quad x = x_1 x_2 \cdots x_n\right\}$$

Comme un sous-groupe de $\mathbb{G}$ contenant F est stable par produit et passage à l'inverse, ce sous-groupe contient évidemment $\mathbb{F}$. On a donc

$$\mathbb{F} \subset \mathrm{Gr}(F)$$

Pour obtenir l'inclusion inverse, il suffit de remarquer que $\mathbb{F}$ est un sous-groupe de $\mathbb{G}$ (il est clairement stable par produit et passage à l'inverse) qui contient F. ■

DÉFINITION 17-1.26 *Une partie F d'un groupe $(\mathbb{G}, .)$ est une partie génératrice de $\mathbb{G}$ (on dit aussi que F est un système de générateurs de $\mathbb{G}$) si et seulement si*

$$\mathrm{Gr}(F) = \mathbb{G}$$

ce qui signifie que tout élément de $\mathbb{G}$ s'écrit comme produit d'éléments de F et F^{-1}.

EXERCICE 17-1.27 Si F engendre $(\mathbb{G}, .)$, et $\varphi : (\mathbb{G}, .) \to (\mathbb{G}', .)$ est un morphisme de groupe, montrer que $\varphi(F)$ engendre $\mathrm{Im}\,\varphi$.

DÉFINITION 17-1.28 *Un groupe $(\mathbb{G}, .)$ est dit monogène s'il est engendré par un de ses éléments :*

$$\exists a \in \mathbb{G} \quad \mathrm{Gr}(a) = \mathbb{G}$$

Comme un produit d'éléments de $\{a, a^{-1}\}$ est une puissance de a, on a alors[6]

$$\mathbb{G} = \{a^n\,,\, n \in \mathbb{Z}\}$$

Un tel groupe est évidemment commutatif.

EXEMPLE 17-1.29 $(\mathbb{Z}, +)$ est un groupe monogène. Montrer qu'il est engendré par $\{1\}$ et aussi par $\{-1\}$.

EXEMPLE 17-1.30 Dans le groupe $(\mathbb{C}^*, \times)$, le sous-groupe monogène engendré par un élément a est l'ensemble des puissances de a. Si a n'est pas une racine de l'unité, il est facile de voir que

$$\forall n, p \in \mathbb{Z} \qquad n \neq p \Rightarrow a^n \neq a^p$$

et l'application

$$(\mathbb{Z}, +) \to (\mathrm{Gr}(a), \times) \qquad n \mapsto a^n$$

[6] En notation additive, on aurait évidemment

$$\mathbb{G} = \{na\,,\, n \in \mathbb{Z}\}$$

est clairement un isomorphisme de groupes : on calcule des produits de puissances de a comme on calcule sur les exposants dans $(\mathbb{Z},+)$. Il n'y a pas de "simplification". Par contre, si a est racine de l'unité avec

$$p = \inf\{n \in \mathbb{N}^* \mid a^n = 1\}$$

il est facile de voir que

$$\mathrm{Gr}(a) = \left\{1, a, a^2, \dots, a^{p-1}\right\}$$

est fini, formé de p éléments distincts 2 à 2. Par exemple $\mathrm{Gr}(i) = \{\pm 1, \pm i\}$. Nous verrons ultérieurement que ces deux situations ne sont pas spécifiques au groupe $(\mathbb{C}^*, \times)$ mais sont tout à fait générales.

EXEMPLE 17-1.31 Soient A, B, C, D les sommets d'un carré du plan euclidien $\mathcal{P}$. On considère le groupe $(\mathbb{G}, \circ)$ des isométries du plan conservant l'ensemble $\{A, B, C, D\}$: on montre que c'est un groupe en prouvant simplement que c'est un sous-groupe du groupe (bien connu) des isométries de $\mathcal{P}$ (s'il n'est pas bien connu, on peut prendre le groupe des bijections affines de $\mathcal{P}$ dans lui-même, ou encore le groupe $(\Sigma_{\mathcal{P}}, \circ)$ des permutations du plan).

Un élément $f \in \mathbb{G}$ conservant $\{A, B, C, D\}$ conserve l'isobarycentre O de ces 4 points, c'est-à-dire le centre du carré $ABCD$. Ramenant l'origine du plan en O, on est alors ramené à un problème vectoriel : déterminer les transformations orthogonales du plan vectoriel P sous-jacent à $\mathcal{P}$ conservant $\left\{\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}, \overrightarrow{OD}\right\}$. Supposons ce plan orienté, et les sommets ordonnés de manière à avoir

$$\left(\widehat{\overrightarrow{OA}, \overrightarrow{OB}}\right) = \left(\widehat{\overrightarrow{OB}, \overrightarrow{OC}}\right) = \left(\widehat{\overrightarrow{OC}, \overrightarrow{OD}}\right) = \left(\widehat{\overrightarrow{OD}, \overrightarrow{OA}}\right) = \frac{\pi}{2} \pmod{2\pi}$$

Le groupe orthogonal $\mathcal{O}(P)$ est composé de rotations et de réflexions (symétries orthogonales par rapport à une droite vectorielle). Il y a évidemment 4 rotations dans le groupe $(\mathbb{G}, \circ)$: si r est la rotation (de centre O) et d'angle $\frac{\pi}{2}$, ce sont

$$r, r^2, r^3 \text{ et } r^4 = id_{\mathcal{P}} = 1_{\mathbb{G}}$$

d'angles respectifs $\frac{\pi}{2}, \pi, \frac{3\pi}{2}$ et 0. Il existe des réflexions dans $\mathbb{G}$. Par exemple la réflexion s par rapport à la droite OA, qui laisse A et C invariants et permute B et D. Comme l'application

$$\mathbb{G} \to \mathbb{G} \qquad f \mapsto s \circ f = sf$$

est une bijection de $\mathbb{G}$ dans lui-même (*cf* proposition **17-1**.3) qui transforme les isométries directes en indirectes et réciproquement, il y a également[7] 4 réflexions dans $\mathbb{G}$

$$sr, sr^2, sr^3 \text{ et } s$$

[7] Plus généralement, s'il existe une isométrie négative conservant une figure, le même raisonnement permet de construire une bijection entre l'ensemble des isométries positives (déplacements) et l'ensemble des isométries négatives (antidéplacements) conservant cette figure.

(s et sr^2 sont les symétries par rapport aux diagonales du carré, sr et sr^3 par rapport aux médianes). On obtient tout élément de $\mathbb{G}$ comme produit formé à l'aide de s et r (on a ici $s = s^{-1}$ et $r^{-1} = r^3$). On a donc

$$\mathbb{G} = \mathrm{Gr}(\{s, r\}) = \left\{id_{\mathcal{P}}, r, r^2, r^3, s, sr, sr^2, sr^3\right\}$$

Il existe un procédé "mécanique" pour ramener un mot construit sur s et r à une de ces 8 écritures "canoniques" : on utilise les relations de "simplification"

$$\begin{cases} s^2 = r^4 = 1_{\mathbb{G}} \\ rs = sr^3 \end{cases}$$

la deuxième relation s'écrivant aussi $srs = r^3$ et se vérifiant par le fait que ces deux rotations envoient A sur D. Elle permet, dans un mot construit sur r et s, de faire disparaître un s qui serait sur la droite d'un produit. On a donc trouvé un système générateurs à deux éléments pour $\mathbb{G}$. On ne peut diminuer ce nombre, puisque $\mathbb{G}$ n'étant pas commutatif n'est évidemment pas monogène.

EXERCICE 17-1.32 Soit $\mathbb{G}$ un groupe fini de cardinal n. Montrer que, si $\mathbf{F}$ est un sous-groupe de $\mathbb{G}$ de cardinal p et $x \in \mathbb{G} - \mathbf{F}$, le sous-groupe de $\mathbb{G}$ engendré par $\{x\} \cup \mathbf{F}$ est de cardinal supérieur ou égal à $2p$. En déduire qu'il existe dans $\mathbb{G}$ un système de générateurs de cardinal inférieur à $\log_2(n)$.

17-1.3 Groupe $(\mathbb{Z}/n\mathbb{Z}, +)$. Groupes cycliques

17-1.3.1 Compatibilité d'une relation d'équivalence et d'une l.c.i.

DÉFINITION 17-1.33 *On considère une loi de composition interne $*$ et une relation d'équivalence $\mathcal{R}$ définies sur un ensemble $\mathbb{X}$. On dit que la loi $*$ est compatible avec $\mathcal{R}$ si et seulement si*

$$\forall x, y, a, b \in \mathbb{X} \quad x\mathcal{R}a \text{ et } y\mathcal{R}b \Rightarrow (x * y)\, \mathcal{R}\, (a * b)$$

En abrégé, on pourra "composer par $*$" des équivalences modulo $\mathcal{R}$ dans $\mathbb{X}$.

Lorsque l'on considère l'ensemble quotient $\mathbb{X}/\mathcal{R}$ formé des classes d'équivalence modulo $\mathcal{R}$, on réalise une partition de $\mathbb{X}$, en regroupant ensemble tous les éléments de $\mathbb{X}$ liés par le relation $\mathcal{R}$. La classe d'équivalence de $a \in \mathbb{X}$ est

$$\overline{a} = \{x \in \mathbb{X} \mid a\mathcal{R}x\}$$

La compatibilité de $*$ avec $\mathcal{R}$ peut alors s'écrire

$$\forall x, y, a, b \in \mathbb{X} \quad \overline{x} = \overline{a} \text{ et } \overline{y} = \overline{b} \Rightarrow \overline{x * y} = \overline{a * b}$$

ce qui montre que la classe modulo $\mathcal{R}$ du composé $x * y$ ne dépend pas vraiment de x et y, mais seulement des classes d'équivalence de x et y. Ceci permet alors de définir une loi de composition interne dans $\mathbb{X}/\mathcal{R}$, appelée **loi quotient de $*$ par $\mathcal{R}$**, que nous noterons encore $*$ s'il n'y a pas d'ambiguïté, par :

$$\overline{x} * \overline{y} \underset{\text{déf}}{=} \overline{x * y}$$

le résultat de l'opération ne dépendant pas des représentants choisis pour $\overline{x}$ et $\overline{y}$. Il importe de bien comprendre que cette définition n'aurait aucun sens s'il n'y avait pas compatibilité entre $*$ et $\mathcal{R}$. La proposition suivante est conséquence immédiate de la définition de la loi quotient :

PROPOSITION 17-1.34 Si la l.c.i $*$ et la relation d'équivalence $\mathcal{R}$ sur l'ensemble $\mathbb{X}$ sont compatibles, la projection canonique

$$p : (\mathbb{X}, *) \to (\mathbb{X}/\mathcal{R}, *) \qquad x \mapsto p(x) = \overline{x}$$

est un morphisme surjectif pour les lois $*$.

Il en résulte aisément que certaines propriétés de $*$ sur $\mathbb{X}$ vont se transporter par p. En particulier :

- Si $*$ est associative dans $\mathbb{X}$, il en est de même de la loi quotient dans $\mathbb{X}/\mathcal{R}$.
- Si $*$ est commutative dans $\mathbb{X}$, il en est de même de la loi quotient dans $\mathbb{X}/\mathcal{R}$.
- Si $\mathbb{X}$ contient un élément neutre e pour $*$, sa classe d'équivalence $\overline{e}$ est élément neutre dans $(\mathbb{X}/\mathcal{R}, *)$.
- Si $\mathbb{X}$ contient un élément neutre e pour $*$ et si $x \in \mathbb{X}$ possède un inverse y à droite (resp. à gauche) pour $*$, alors $\overline{y}$ est inverse à droite (resp. à gauche) de $\overline{x}$ dans $(\mathbb{X}/\mathcal{R}, *)$.

COROLLAIRE 17-1.35 Sous les hypothèses précédentes, si $(\mathbb{X}, *)$ est un groupe, il en est de même de $(\mathbb{X}/\mathcal{R}, *)$.

EXERCICE 17-1.36 On dit que la l.c.i $*$ et la relation d'équivalence $\mathcal{R}$ sur $\mathbb{X}$ sont compatibles à gauche ssi

$$\forall x, y, a \in \mathbb{X} \qquad x\mathcal{R}y \Rightarrow (a * x)\,\mathcal{R}\,(a * y)$$

et on définit de même la compatibilité à droite (distinction sans intérêt si $*$ est commutative). Montrer que $*$ et $\mathcal{R}$ sont compatibles (au sens de la définition **17-1**.33) si et seulement si elles sont compatibles à la fois à droite et à gauche.

EXERCICE 17-1.37 Soit $(\mathbb{G}, .)$ un groupe, et $\mathcal{R}$ une relation d'équivalence dans $\mathbb{G}$. Montrer que, si $\mathcal{R}$ est compatible à gauche avec ".", alors

$$\forall x, y \in \mathbb{G} \quad x\mathcal{R}y \Leftrightarrow x^{-1}y \in \overline{1}_{\mathbb{G}}$$

et montrer que la classe $\overline{1}_{\mathbb{G}}$ de l'élément neutre est un sous-groupe de $\mathbb{G}$. Montrer que, réciproquement, si $\mathbb{H}$ est un sous-groupe de $\mathbb{G}$, la relation binaire dans $\mathbb{G}$ définie par

$$x\mathcal{R}y \Leftrightarrow x^{-1}y \in \mathbb{H}$$

est une relation d'équivalence compatible à gauche avec ".", pour laquelle $\mathbb{H}$ est exactement la classe d'équivalence de $1_{\mathbb{G}}$, et vérifiant

$$\forall x \in \mathbb{G} \quad \overline{x} = x.\mathbb{H}$$

qu'on appelle classe de x à gauche modulo $\mathbb{H}$.

Lorsque le groupe $(\mathbb{G}, .)$ est commutatif, ces résultats caractérisent les relations d'équivalence dans $\mathbb{G}$ compatibles avec la loi du groupe. Dans $(\mathbb{Z}, +)$, nous étudierons dans la section suivante la congruence modulo un entier n qui est une relation de ce type.

EXERCICE 17-1.38 (Suite de l'exercice précédent) Si le groupe $(\mathbb{G},.)$ n'est pas commutatif, et $\mathbb{H}$ est un sous-groupe de $\mathbb{G}$, montrer que la relation

$$x\mathcal{R}y \Leftrightarrow x^{-1}y \in \mathbb{H}$$

est aussi compatible à droite avec "." si et seulement si

$$\forall x \in \mathbb{G} \quad x.\mathbb{H} = \mathbb{H}.x$$

ce qui revient aussi à dire que

$$\forall x \in \mathbb{G} \quad x^{-1}\mathbb{H}x = \mathbb{H}$$

Nous avons vu (*cf* exercice **17-1.22**) que le noyau d'un morphisme de groupes $\varphi : (\mathbb{G},.) \to (\mathbb{G}',.)$ vérifie toujours cette propriété : la relation d'équivalence associée à $\ker\varphi$ est alors

$$x\mathcal{R}y \Leftrightarrow x^{-1}y \in \ker\varphi \Leftrightarrow \varphi(x) = \varphi(y)$$

et on vérifie aisément qu'elle est compatible à la fois à droite et à gauche avec la loi de $\mathbb{G}$.

17-1.3.2 Congruence modulo n

DÉFINITION 17-1.39 *Soit $n \in \mathbb{Z}$. On appelle congruence modulo n la relation binaire définie sur $\mathbb{Z}$ par*

$$x \equiv y \mod n \Leftrightarrow x - y \in n\mathbb{Z} \Leftrightarrow n \text{ divise } (x-y)$$

$x \equiv y \mod n$ se lira "x est congru à y modulo n".

Lorsque $n = 0$, on obtient la relation d'égalité. Pour $n = 1$, c'est la relation triviale qui relie tous les éléments de $\mathbb{Z}$. Comme de plus la congruence modulo n est identique à la congruence modulo $-n$, on supposera dans la suite $n \geqslant 2$.

THÉORÈME 17-1.40 La congruence modulo n est une relation d'équivalence dans $\mathbb{Z}$. Tout entier relatif est congru modulo n à son reste dans la division euclidienne par n. Chaque classe d'équivalence possède un unique représentant dans $\{0, 1, \ldots, n-1\}$. L'ensemble de ces classes d'équivalence, noté $\mathbb{Z}/n\mathbb{Z}$, est donc de cardinal n. On a

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

en notant, pour $x \in \mathbb{Z}$

$$\overline{x} = x + n\mathbb{Z}$$

la classe de x pour cette relation de congruence.

Démonstration : Le fait que l'on ait une relation d'équivalence se vérifie aisément. Si x est un entier quelconque, la division euclidienne de x par n s'écrit

$$x = nq + r \text{ avec } 0 \leqslant r \leqslant n-1$$

et comme $n|(x-r)$, on a bien $\overline{x}=\overline{r}$. De plus

$$0 \leqslant r_1 < r_2 \leqslant n-1 \Longrightarrow 0 < r_2 - r_1 < n$$

ce qui montre que n ne peut diviser $r_2 - r_1$, et que par conséquent $\overline{r_2} \neq \overline{r_1}$. ■

On vérifie facilement que l'addition dans $\mathbb{Z}$ est compatible avec cette relation de congruence. Il en résulte que l'on peut définir sur $\mathbb{Z}/n\mathbb{Z}$ une addition par

$$\overline{x}+\overline{y}=\overline{x+y}$$

La projection canonique $x \mapsto \overline{x}$ étant un morphisme surjectif, la structure de groupe commutatif de $(\mathbb{Z},+)$ se transporte, et on obtient :

THÉORÈME 17-1.41 $(\mathbb{Z}/n\mathbb{Z},+)$ **est un groupe commutatif monogène.**

Démonstration : On applique le corollaire **17-1**.35. De plus, il est clair que $\{\overline{1}\}$ engendre $(\mathbb{Z}/n\mathbb{Z},+)$, puisque pour $k \in \{1,\dots,n-1\}$, on a simplement

$$\overline{k}=\overline{\underbrace{1+\cdots+1}_{k \text{ fois}}}=\underbrace{\overline{1}+\cdots+\overline{1}}_{k \text{ fois}}=k.\overline{1}$$

(attention : la notation est ici additive). Il s'agit en fait d'une application directe du résultat de l'exercice **17-1**.27. ■

Si l'on décide de représenter chaque classe de congruence modulo n par son représentant "canonique" dans $\{0,\dots,n-1\}$, on calcule sur l'addition dans $\mathbb{Z}/n\mathbb{Z}$ tout simplement en additionnant les représentants, et en conservant du résultat le reste dans sa division euclidienne par n. Par exemple la table d'addition de $\mathbb{Z}/5\mathbb{Z}$ est

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$

17-1.3.3 Structure d'un groupe monogène. Ordre d'un élément d'un groupe

Soit $(\mathbb{G},.)$ un groupe et $a \in \mathbb{G}$. Nous savons que

$$\mathrm{Gr}(a)=\{a^n,\ n\in\mathbb{Z}\}$$

L'application $\varphi : (\mathbb{Z},+) \to (\mathrm{Gr}(a),.)$ définie par $\varphi(n)=a^n$ est clairement un morphisme surjectif de groupes. Deux situations sont alors envisageables :

- φ est un isomorphisme de groupes, ce qui revient à supposer que l'on a $\ker\varphi = \{0\}$, c'est-à-dire que

$$\forall k \in \mathbb{Z} \quad a^k = 1_G \Leftrightarrow k = 0$$

Le groupe $(\mathrm{Gr}(a), .)$ est infini, isomorphe à $(\mathbb{Z}, +)$; deux exposants distincts donnent des puissances de a distinctes. On dit alors que

$$a \text{ est un élément d'ordre infini de } (\mathbb{G}, .)$$

- φ n'est pas injectif. Son noyau est un sous-groupe de $(\mathbb{Z}, +)$ différent de $\{0\}$, donc

$$\ker\varphi = m\mathbb{Z}$$

où $m \in \mathbb{N}^*$ (le cas $m = 1$ correspond évidemment à $a = 1_G$). On a alors

$$\forall k \in \mathbb{Z} \quad a^k = 1_G \Leftrightarrow m \text{ divise } k$$

soit encore, puisque pour $k, l \in \mathbb{Z}$ on a $\varphi(k) = \varphi(l) \Leftrightarrow k - l \in \ker\varphi$:

$$\forall k, l \in \mathbb{Z} \quad a^k = a^l \Leftrightarrow k \equiv l \bmod m \qquad (*)$$

L'entier naturel m est **le plus petit entier > 0 vérifiant** $a^m = 1_G$: "plus petit" évidemment au sens de le relation d'ordre naturelle sur $\mathbb{N}^*$, mais plus précisément **diviseur de tout entier** k vérifiant $a^k = 1_G$. On dit alors que

$$m \text{ est l'ordre de } a \text{ dans le groupe } (\mathbb{G}, .)$$

La propriété $(*)$ montre alors qu'il y a autant d'éléments dans $\mathrm{Gr}(a)$ que de classes de congruence modulo m dans $\mathbb{Z}$. Le groupe engendré par a est donc de cardinal m, avec

$$\mathrm{Gr}(a) = \{1_G, a, a^2, \ldots, a^{m-1}\}$$

Plus précisément, on peut définir un isomorphisme de groupes

$$\begin{aligned} \widetilde{\varphi} : (\mathbb{Z}/m\mathbb{Z}, +) &\to (\mathrm{Gr}(a), .) \\ \overline{k} &\mapsto a^k \end{aligned}$$

(la propriété $(*)$ dit simplement que cette application est bien définie et injective; le fait que $\widetilde{\varphi}$ soit un morphisme surjectif est alors évident). Le groupe monogène $\mathrm{Gr}(a)$, fini de cardinal m est alors appelé **groupe cyclique d'ordre** m. Le terme "cyclique" traduit simplement le fait que la suite $(a^k)_{k\in\mathbb{N}}$ est périodique, de plus petite période égale à m.

Résumons la discussion précédente :

THÉORÈME 17-1.42 (ET DÉFINITION) **Soit a un élément d'un groupe $(\mathbb{G},.)$. On dit que a est d'ordre fini si et seulement si $\mathrm{Gr}(a)$ est fini. Ceci équivaut à**

$$\exists k \in \mathbb{N}^* \quad a^k = 1_{\mathbb{G}}$$

On appelle ordre de a et on note $[a]$ le plus petit entier k vérifiant cette égalité. On a alors

$$\forall k \in \mathbb{Z} \quad a^k = 1_{\mathbb{G}} \Leftrightarrow [a] \text{ divise } k$$

et $\mathrm{Gr}(a)$ est de cardinal $[a]$, avec

$$\mathrm{Gr}(a) = \left\{1_{\mathbb{G}}, a, a^2, \dots, a^{[a]-1}\right\}$$

L'application $\varphi : (\mathbb{Z}/[a]\mathbb{Z}, +) \to (\mathrm{Gr}(a), .)$ définie par $\overline{k} \mapsto a^k$ est un isomorphisme de groupes.

Si a est d'ordre infini, l'application $\psi : (\mathbb{Z}, +) \to (\mathrm{Gr}(a), .)$ définie par $\psi(k) = a^k$ est un isomorphisme.

Il est donc clair que, **dans un groupe fini, tout élément est d'ordre fini.**

DÉFINITION 17-1.43 *Si $(\mathbb{G},.)$ est un groupe fini, on appelle ordre de $\mathbb{G}$, et on note $[\mathbb{G}]$ le cardinal de $\mathbb{G}$.*

DÉFINITION 17-1.44 *Un groupe $(\mathbb{G},.)$ est dit cyclique s'il est monogène fini : il existe $a \in \mathbb{G}$ avec $[a] = [\mathbb{G}]$. Le groupe $(\mathbb{G},.)$ est alors isomorphe à $(\mathbb{Z}/[\mathbb{G}]\mathbb{Z}, +)$*

EXERCICE 17-1.45 Si m est un entier $\geqslant 2$, montrer que les groupes d'ordre m^2

$$(\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}, +) \text{ et } (\mathbb{Z}/m^2\mathbb{Z}, +)$$

ne sont pas isomorphes.

EXERCICE 17-1.46 Soient a et b deux entiers $\geqslant 2$ premiers entre eux. Si k est un entier, on note $\overline{k}$ (resp. $\tilde{k}$) sa classe de congruence modulo a (resp. b). Déterminer $\left[\left(\overline{1}, \tilde{1}\right)\right]$ dans le groupe $(\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}, +)$. En déduire que ce groupe est cyclique.

EXERCICE 17-1.47 Soit $(\mathbb{G},.)$ un groupe *commutatif* fini. Montrer que, si x et $y \in \mathbb{G}$,

$$[x] \wedge [y] = 1 \Rightarrow [xy] = [x][y]$$

(Indication : si $k \in \mathbb{Z}$ vérifie $(xy)^k = 1_{\mathbb{G}}$, calculer $(xy)^{k[x]}$). Généraliser au produit de p éléments de $\mathbb{G}$ dont les ordres sont premiers entre eux deux à deux.

EXERCICE 17-1.48 (suite de l'exercice précédent) : soit $(\mathbb{G},.)$ un groupe *commutatif* fini. On note m le P.P.C.M des ordres de tous les éléments de $\mathbb{G}$. On considère la décomposition de m en facteurs premiers

$$m = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$$

Montrer qu'il existe dans $\mathbb{G}$ un élément x_1 dont l'ordre s'écrit

$$[x_1] = p_1^{\alpha_1}.d_1 \text{ avec } d_1 \wedge p_1 = 1$$

On pose $y_1 = x_1^{d_1}$. Quel est l'ordre de y_1 ? Montrer qu'il existe dans $\mathbb{G}$ un élément d'ordre m (dont l'ordre est donc multiple des ordres de tous les éléments de $\mathbb{G}$).

EXERCICE 17-1.49 Soit $(\mathbb{G},.)$ un groupe fini vérifiant

$$\forall x \in \mathbb{G} \quad x^2 = 1_{\mathbb{G}}$$

Montrer que $\mathbb{G}$ est commutatif. Si $\mathbb{F}$ est un sous-groupe de $\mathbb{G}$ et $x \in \mathbb{G} - \mathbb{F}$, montrer que $\mathbb{F} \cup x\mathbb{F}$ est un sous-groupe de $\mathbb{G}$. En déduire qu'il existe $p \in \mathbb{N}$ avec

$$[\mathbb{G}] = 2^p$$

et montrer que, plus précisément, $\mathbb{G}$ est isomorphe à $((\mathbb{Z}/2\mathbb{Z})^p, +)$.

EXERCICE 17-1.50 Soit $\mathbb{X}$ un sous-ensemble fini de $\mathcal{GL}_n(\mathbb{C})$ stable par multiplication. Montrer que toute matrice de $\mathbb{X}$ est diagonalisable.

17-1.3.4 Générateurs d'un groupe cyclique

THÉORÈME 17-1.51 Soit $\mathbb{G} = \mathrm{Gr}(a)$ un groupe cyclique d'ordre m

$$\mathbb{G} = \left\{1_{\mathbb{G}}, a, a^2, \ldots, a^{m-1}\right\}$$

Si $k \in \mathbb{Z}$, a^k engendre $\mathbb{G}$ si et seulement si k est premier avec m.

Démonstration : Considérons $\mathrm{Gr}\left(a^k\right)$ le sous-groupe de $\mathbb{G}$ engendré par a^k. Il est facile de voir que ce sous-groupe est égal à $\mathbb{G}$ si et seulement si

$$a \in \mathrm{Gr}\left(a^k\right)$$

ce qui s'écrit également

$$\exists l \in \mathbb{Z} \quad \left(a^k\right)^l = a$$

D'après ce qui a été vu dans la section précédente, cela signifie que l'ordre de a divise $kl - 1$:

$$\exists l \in \mathbb{Z} \quad \exists n \in \mathbb{Z} \quad kl + nm = 1$$

Il s'agit exactement de l'identité de Bézout, traduisant le fait que m et k sont premiers entre eux. ∎

Avec une notation additive, et puisque dans $\mathbb{Z}/m\,\mathbb{Z}$ on a $\overline{k} = k.\overline{1}$, on obtient :

COROLLAIRE 17-1.52 Si $k \in \mathbb{Z}$, la classe $\overline{k}$ de k modulo m engendre le groupe $(\mathbb{Z}/m\,\mathbb{Z}, +)$ si et seulement si

$$k \wedge m = 1$$

EXEMPLE 17-1.53 Pour n entier $\geqslant 2$, l'ensemble des racines complexes $n^{\text{ièmes}}$ de l'unité

$$\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$$

est un sous-groupe de $(\mathbb{C}^*, \times)$. Il est clairement cyclique d'ordre n engendré par exemple par

$$\omega = e^{\frac{2i\pi}{n}}$$

Un générateur quelconque du groupe $\mathbb{U}_n$ est appelé racine **primitive** $n^{\text{ième}}$ de l'unité. Il s'agit donc d'un complexe de la forme

$$\omega^k = e^{\frac{2ik\pi}{n}}$$

avec k entier premier avec n.

17-2 Groupe opérant sur un ensemble

Dans cette section, nous considérerons un élément d'un groupe comme une transformation "géométrique" d'un ensemble.

17-2.1 Définition

Dans ce qui suit, $\mathbb{X}$ est un ensemble (non vide) quelconque. Nous appellerons parfois "point" un élément quelconque de $\mathbb{X}$, par référence à la géométrie élémentaire, mais les éléments de $\mathbb{X}$ peuvent être de natures très diverses.

DÉFINITION 17-2.1 *Une opération d'un groupe* $(\mathbb{G},.)$ *sur* $\mathbb{X}$ *est une application*

$$\varphi : \mathbb{G} \times \mathbb{X} \to \mathbb{X} \qquad (\sigma, x) \mapsto \varphi(\sigma, x)$$

vérifiant :

$$\forall x \in \mathbb{X} \quad \varphi(1_{\mathbb{G}}, x) = x$$
$$\forall \sigma, \sigma' \in \mathbb{G} \quad \varphi(\sigma, \varphi(\sigma', x)) = \varphi(\sigma\sigma', x)$$

Notation : Les écritures précédentes étant peu commodes, on utilisera plutôt, lorsqu'il n'y a pas d'ambiguïté, une notation du type "opération externe" sur $\mathbb{X}$:

$$\varphi(\sigma, x) = \sigma.x$$

est le point de $\mathbb{X}$ résultant de l'opération de $\sigma \in \mathbb{G}$ sur $x \in \mathbb{X}$. Il faut alors être vigilant puisque nous notons par "." aussi bien la loi interne dans le groupe $\mathbb{G}$ que l'opération de $\mathbb{G}$ sur $\mathbb{X}$. Les propriétés de la définition précédente s'écrivent ainsi

$$\forall x \in \mathbb{X} \quad 1_{\mathbb{G}}.x = x$$
$$\forall \sigma, \sigma' \in \mathbb{G} \quad \forall x \in \mathbb{X} \quad \sigma.(\sigma'.x) = (\sigma\sigma').x$$

EXEMPLE 17-2.2 Si $\mathbb{X}$ est un ensemble quelconque, le groupe $(\Sigma_{\mathbb{X}}, \circ)$ des permutations de $\mathbb{X}$ opère de manière naturelle sur $\mathbb{X}$ par $(f, x) \mapsto f.x = f(x)$.

EXEMPLE 17-2.3 Si $\mathcal{A}$ est un espace affine basé sur un espace vectoriel $\mathbb{E}$, le groupe additif $(\mathbb{E}, +)$ opère sur $\mathcal{A}$ par l'intermédiaire des translations :

$$(\vec{x}, A) \mapsto \vec{x}.A = A + \vec{x}$$

(et bien évidemment, dans ce cas, on n'utiliserons pas la notation $\vec{x}.A$).

EXEMPLE 17-2.4 Si $\mathbb{E}$ est un espace vectoriel, on peut faire opérer naturellement le groupe linéaire $(\mathcal{GL}_{\mathbb{K}}(\mathbb{E}), \circ)$ sur l'espace vectoriel lui-même, ou sur l'ensemble $\mathbb{X}$ des sous-espaces vectoriels de $\mathbb{E}$ par

$$\mathcal{GL}_{\mathbb{K}}(\mathbb{E}) \times \mathbb{E} \quad (u, x) \mapsto u(x)$$
$$\mathcal{GL}_{\mathbb{K}}(\mathbb{E}) \times \mathbb{X} \quad (u, \mathbb{F}) \mapsto u(\mathbb{F})$$

Dans chacun des exemples précédents, on fait opérer des bijections sur l'ensemble $\mathbb{X}$. Cette situation est générale :

THÉORÈME 17-2.5 Si $(\sigma, x) \mapsto \sigma.x$ est une opération d'un groupe $(\mathbb{G}, .)$ sur un ensemble $\mathbb{X}$, pour tout $\sigma \in \mathbb{G}$ l'application τ_σ définie par

$$\tau_\sigma : \mathbb{X} \to \mathbb{X} \quad x \mapsto \tau_\sigma(x) = \sigma.x$$

est une bijection de $\mathbb{X}$ dans lui-même, et l'application

$$\psi : (\mathbb{G}, .) \to (\Sigma_{\mathbb{X}}, \circ) \quad \sigma \mapsto \tau_\sigma$$

est un morphisme de groupes.

Démonstration : Soit $y \in \mathbb{X}$. Montrons que l'équation $\sigma.x = y$ possède une unique solution $x \in \mathbb{X}$. Si x existe, on a nécessairement

$$\sigma^{-1}.y = \sigma^{-1}.(\sigma.x) = \left(\sigma^{-1}\sigma\right).x = 1_{\mathbb{G}}.x = x$$

d'après les propriétés de l'opération de $\mathbb{G}$ sur $\mathbb{X}$. On vérifie ensuite aisément que x ainsi défini est bien solution de l'équation. De plus la propriété

$$\forall \sigma, \sigma' \in \mathbb{G} \quad \forall x \in \mathbb{X} \quad \sigma.(\sigma'.x) = (\sigma\sigma').x$$

s'écrit exactement $\tau_\sigma \circ \tau_{\sigma'} = \tau_{\sigma\sigma'}$, ce qui montre bien que ψ est un morphisme de groupes. ■

EXERCICE 17-2.6 Montrer que réciproquement, tout morphisme de groupes

$$\psi : (\mathbb{G}, .) \to (\Sigma_{\mathbb{X}}, \circ)$$

permet de définir une opération de $\mathbb{G}$ sur $\mathbb{X}$ par $\sigma.x = \psi(\sigma)(x)$.

En conclusion, considérer une opération de $\mathbb{G}$ sur $\mathbb{X}$, c'est associer à tout élément de $\mathbb{G}$ une permutation de $\mathbb{X}$, la composition de ces permutations correspondant à la loi de composition interne dans $\mathbb{G}$. Tout élément de $\mathbb{G}$ est alors considéré comme une transformation "géométrique" de $\mathbb{X}$.

17-2.2 Opérations d'un groupe sur lui-même

Nous étudions ici deux exemples classiques d'opérations d'un groupe sur lui-même, dont nous verrons quelques applications remarquables ultérieurement.

17-2.2.1 Translations à gauche

Si $(\mathbb{G}, .)$ est un groupe, on peut le considérer également comme ensemble de "points", et faire opérer $\mathbb{G}$ sur lui-même par translation à gauche :

$$(x, y) \in \mathbb{G} \times \mathbb{G} \mapsto x.y = xy$$

On montre aisément que les propriétés de définition sont vérifiées. Si on note τ_x la translation à gauche par x définie sur $\mathbb{G}$, le théorème **17-2.5** montre alors que l'application

$$\psi : (\mathbb{G}, .) \to (\Sigma_{\mathbb{G}}, \circ) \quad x \mapsto \tau_x$$

est un morphisme de groupes. Ce morphisme est ici clairement injectif, puisque

$$\forall x \in \mathbb{G} \quad \tau_x = id_{\mathbb{G}} \Leftrightarrow x = 1_{\mathbb{G}}$$

On en déduit que ψ induit un isomorphisme entre $\mathbb{G}$ et $\psi(\mathbb{G})$, sous-groupe de $(\Sigma_{\mathbb{G}}, \circ)$. Comme la structure de $(\Sigma_{\mathbb{X}}, \circ)$ ne dépend que du "cardinal" de $\mathbb{X}$ (voir exemple **17-1**.13), nous obtenons le théorème (Cayley) :

THÉORÈME 17-2.7 Si $(\mathbb{G}, .)$ **est un groupe fini de cardinal** n**, il est isomorphe à un sous-groupe du groupe symétrique** $(S_n, \circ)$**.**

REMARQUE 17-2.8 Si $\mathbb{H}$ est un sous-groupe de $\mathbb{G}$, on peut aussi faire opérer $\mathbb{H}$ sur $\mathbb{G}$ par translation à gauche (en réduisant le domaine des opérateurs). Cette remarque est valable pour toute action d'un groupe $\mathbb{G}$ sur un ensemble $\mathbb{X}$.

17-2.2.2 Automorphismes intérieurs

Il s'agit d'une notion qui n'a d'intérêt que dans un groupe non commutatif.

DÉFINITION 17-2.9 *Si x est un élément d'un groupe* $(\mathbb{G}, .)$, *on appelle automorphisme intérieur associé à x l'application*

$$i_x : \mathbb{G} \to \mathbb{G} \quad y \mapsto xyx^{-1}$$

Il s'agit bien d'un automorphisme du groupe $\mathbb{G}$, puisque pour $y, z \in \mathbb{G}$

$$i_x(y)\, i_x(z) = \left(xyx^{-1}\right)\left(xzx^{-1}\right) = xyzx^{-1} = i_x(yz)$$

et on a clairement

$$\forall y, z \in \mathbb{G} \quad z = x^{-1}yx \Leftrightarrow y = xzx^{-1}$$

ce qui donne

$$(i_x)^{-1} = i_{x^{-1}}$$

On a de même

$$\forall x, y \in \mathbb{G} \quad i_x \circ i_y = i_{xy}$$

et l'application $x \mapsto i_x$ est un morphisme de $(\mathbb{G}, .)$ dans $(\mathrm{Aut}(\mathbb{G}), \circ)$. On vérifie alors facilement (*cf* exercice **17-2**.6) que

$$x \bullet y = xyx^{-1}$$

définit une opération[8] de $\mathbb{G}$ sur lui-même.

REMARQUE 17-2.10 Comme pour $x \in \mathbb{G}$, i_x est un automorphisme de $\mathbb{G}$, si $\mathbb{H}$ est un sous-groupe de $\mathbb{G}$, il en est de même de $i_x(\mathbb{H}) = x\mathbb{H}x^{-1}$. On peut donc faire opérer $\mathbb{G}$ sur l'ensemble de ses sous-groupes par automorphismes intérieurs. Cela n'a pas de sens avec les translations à gauche.

[8]On notera que la similitude associée à une matrice carrée inversible $P \in \mathcal{GL}_n(\mathbb{K})$ est un automorphisme intérieur de $(\mathcal{GL}_n(\mathbb{K}), \circ)$.

17-2.3 Orbite d'un point

17-2.3.1 Orbite d'un élément

THÉORÈME 17-2.11 (ET DÉFINITION) Si $(\mathbb{G},.)$ est un groupe opérant sur un ensemble $\mathbb{X}$, la relation binaire sur $\mathbb{X}$ définie par

$$x\mathcal{R}y \Leftrightarrow \exists\sigma\in\mathbb{G}\quad y=\sigma.x$$

est une relation d'équivalence. La classe d'équivalence d'un point $x\in\mathbb{X}$

$$\mathcal{O}(x)=\{\sigma.x,\ \sigma\in\mathbb{G}\}$$

est appelée orbite de x (pour l'opération de $\mathbb{G}$ sur $\mathbb{X}$ envisagée). Les différentes orbites réalisent donc une partition de $\mathbb{X}$.

Démonstration : Exercice.

EXEMPLE 17-2.12 Soit $\mathcal{A}$ un espace affine de dimension 3 basé sur un espace vectoriel $\mathbb{E}$. Si on fait opérer les vecteurs de $\mathbb{E}$ sur les points de $\mathcal{A}$ par translation, il y a une seule orbite qui est l'espace $\mathcal{A}$ entier. Si on fait opérer un sous-espace vectoriel $\mathbb{F}$ de $\mathbb{E}$ sur $\mathcal{A}$, l'orbite d'un point $A\in\mathcal{A}$ est $A+\mathbb{F}$, sous-espace affine de direction $\mathbb{F}$ passant par A.

EXERCICE 17-2.13 Soit $\mathbb{E}_n$ un espace vectoriel de dimension n. On fait opérer de manière naturelle le groupe $\mathcal{GL}(\mathbb{E}_n)$ sur l'ensemble $\mathbf{X}$ des sous-espaces vectoriels de $\mathbb{E}$. Quelle est l'orbite d'un s.e.v. donné de $\mathbb{E}_n$. Combien y a-t-il d'orbites distinctes ?

EXERCICE 17-2.14 Soit $\mathbb{E}_n$ un espace vectoriel euclidien. Le groupe orthogonal opère naturellement sur $\mathbb{E}_n$. Décrire l'orbite d'un vecteur $x\in\mathbb{E}_n$.

Une opération d'un groupe sur un ensemble permet donc d'obtenir une partition "géométrique" de cet ensemble. Comme nous allons le voir dans les sections qui suivent, ce résultat, banal en apparence, s'avère très efficace.

17-2.3.2 Théorème de Lagrange

Soit $(\mathbb{G},.)$ un groupe et $\mathbb{H}$ un sous-groupe de $\mathbb{G}$. On fait ici opérer $\mathbb{H}$ sur $\mathbb{G}$ par translation à gauche. L'orbite d'un élément $x\in\mathbb{G}$ est alors

$$\mathbb{H}.x=\{h.x,\quad h\in\mathbb{H}\}$$

qu'on appelle classe de x à droite modulo $\mathbb{H}$ (*cf* exercices **17-1**.37 et **17-1**.38). Cette classe est évidemment équipotente à $\mathbb{H}$ (puisque la translation à droite par x est une bijection de $\mathbb{G}$ dans lui-même). Deux orbites sont soit distinctes, soit confondues : pour x et $y\in\mathbb{G}$

$$\mathbb{H}.x=\mathbb{H}.y\quad\text{ou}\quad\mathbb{H}.x\cap\mathbb{H}.y=\emptyset$$

(selon que xy^{-1} est ou n'est pas dans $\mathbb{H}$ comme on le vérifie aisément).

Si $\mathbb{G}$ est fini et m est le nombre (fini !) d'orbites distinctes, comme chaque orbite a pour cardinal l'ordre de $\mathbb{H}$, nous avons donc

$$[\mathbb{G}]=m\,[\mathbb{H}]$$

Ceci démontre le théorème de Lagrange :

THÉORÈME 17-2.15 Si $\mathbb{G}$ est un groupe fini, l'ordre de tout sous-groupe de $\mathbb{G}$ est un diviseur de l'ordre de $\mathbb{G}$.

Si $a \in \mathbb{G}$, l'ordre $[a]$ de a est également l'ordre du sous-groupe (cyclique) $\mathrm{Gr}(a)$ engendré par a. On a donc :

COROLLAIRE 17-2.16 Dans un groupe fini, l'ordre de tout élément est un diviseur de l'ordre du groupe.

COROLLAIRE 17-2.17 Si p est un nombre premier, tout groupe fini de cardinal p est cyclique.

Démonstration : Si $\mathbb{G}$ est un groupe d'ordre p, tout élément de $\mathbb{G}$ distinct de $1_\mathbb{G}$ a pour ordre un diviseur de p différent de 1. Ce ne peut être que p. En conséquence, pour $x \in \mathbb{G} - \{1_\mathbb{G}\}$, nous avons

$$[\mathrm{Gr}(x)] = [\mathbb{G}]$$

et donc $\mathbb{G} = \mathrm{Gr}(x)$ est cyclique. ■

Il n'y a donc, à isomorphisme près, qu'une seule structure de groupe fini d'ordre premier p. C'est celle de $(\mathbb{Z}/p\mathbb{Z}, +)$. Un tel groupe est donc en particulier nécessairement commutatif.

REMARQUE 17-2.18 Si $\mathbb{G}$ est un groupe fini et $\mathbb{H}$ est un sous-groupe de $\mathbb{G}$, le quotient $\frac{[\mathbb{G}]}{[\mathbb{H}]}$ est, si l'on reprend la démonstration du théorème de Lagrange, égal au nombre de classes à droite distinctes modulo $\mathbb{H}$. C'est aussi le nombre de classes à gauche modulo $\mathbb{H}$ (*cf* exercice **17-1**.37) où l'on voit que ces classes sont des classes d'équivalence pour un relation dans $\mathbb{G}$: deux classes sont donc distinctes ou confondues). On note $[\mathbb{G} : \mathbb{H}]$ ce rapport, qu'on appelle indice de $\mathbb{H}$ dans le groupe $\mathbb{G}$:

$$[\mathbb{G} : \mathbb{H}] = \frac{[\mathbb{G}]}{[\mathbb{H}]}$$

Lorsque $\mathbb{G}$ n'est pas fini, on dira de même qu'un sous-groupe $\mathbb{H} \subset \mathbb{G}$ est d'indice fini[9] dans $\mathbb{G}$ si le nombre m de classes à droite modulo $\mathbb{H}$ est fini, et m est alors, par définition l'indice $[\mathbb{G} : \mathbb{H}]$ de $\mathbb{G}$ dans $\mathbb{H}$. En utilisant l'application $x \mapsto x^{-1}$, montrer que c'est aussi le nombre de classes à gauche modulo $\mathbb{H}$.

EXERCICE 17-2.19 Soit $\mathbb{G} = \mathrm{Gr}(a)$ un groupe cyclique d'ordre n. Montrer que tout sous-groupe de $\mathbb{G}$ est cyclique (on peut considérer le morphisme $k \mapsto a^k$ de $\mathbb{Z}$ dans $\mathbb{G}$). Plus précisément, montrer que, pour tout diviseur d de n, $\mathbb{G}$ possède un unique sous-groupe (cyclique) d'ordre d.

EXERCICE 17-2.20 Soit $\mathbb{G}$ et $\mathbb{H}$ deux groupes, et $\varphi : \mathbb{G} \to \mathbb{H}$ un morphisme. Montrer que, si $\mathbb{G}$ est fini,

$$[\mathbb{G}] = [\ker\varphi] \times [\mathrm{Im}\,\varphi]$$

---

[9] Par exemple, si $\mathbb{E}$ est un espace vectoriel euclidien, on a

$$[\mathcal{O}(\mathbb{E}), \mathcal{O}^+(\mathbb{E})] = 2$$

puisque, si r est une réflexion quelconque, $\mathcal{O}(\mathbb{E}) = \mathcal{O}^+(\mathbb{E}) \cup \mathcal{O}^+(\mathbb{E}).r$, l'ensemble $\mathcal{O}^+(\mathbb{E}).r$ contenant toutes les transformations orthogonales indirectes de $\mathbb{E}$.

17-2.3.3 Décomposition d'une permutation en produit de cycles "disjoints"

On travaille ici dans le groupe $(S_n, \circ)$ des permutations de $\{1, \ldots, n\}$ avec $n \geqslant 2$.

DÉFINITION 17-2.21 *Si k est un entier vérifiant $2 \leqslant k \leqslant n$, et $S \subset \{1, \ldots, n\}$ est une partie de cardinal k, une permutation $c \in S_n$ est dite cycle de longueur k opérant sur S si et seulement s'il existe une numérotation des éléments de S, telle que $S = \{a_1, a_2, \ldots, a_k\}$, vérifiant*

$$\forall i \in \{1, \ldots, n\} - S \qquad c(i) = i$$

$$\forall j \in \{1, \ldots, k-1\} \quad c(a_j) = a_{j+1} \quad \text{et} \quad c(a_k) = a_1$$

On notera en abrégé

$$c = (a_1, a_2, \ldots, a_k)$$

cette permutation "circulaire" opérant sur S. Une **transposition** $\tau = (i, j)$ pour $i \neq j$ dans $\{1, \ldots, n\}$ est un cycle de longueur 2. Elle vérifie $\tau(i) = j$, $\tau(j) = i$ et $\tau(k) = k$ pour $k \notin \{i, j\}$. Il est facile de voir qu'un cycle de longueur k est d'ordre k dans le groupe S_n.

EXERCICE 17-2.22 Si S est une partie de $\{1, \ldots, n\}$ de cardinal k, montrer qu'il y a $(k-1)!$ cycles de longueur k distincts opérant sur S.

En utilisant la notion d'orbite, nous allons montrer que les cycles engendrent le groupe S_n, et plus précisément :

THÉORÈME 17-2.23 Si σ est un élément de S_n différent de l'identité, il se décompose comme produit (commutatif) de cycles opérant sur des ensembles disjoints deux à deux. Cette décomposition est unique à l'ordre des facteurs près.

Démonstration : Soit $p = [\sigma]$ l'ordre de la permutation σ. Le groupe

$$\mathbb{G} = \mathrm{Gr}(\sigma) = \left\{id, \sigma, \ldots, \sigma^{p-1}\right\}$$

opère de manière naturelle sur $\{1, \ldots, n\}$ par

$$\left(\sigma^k, y\right) \mapsto \sigma^k(y)$$

Soit $x \in \{1, \ldots, n\}$, et $\mathcal{O}(x)$ son orbite pour cette action

$$\mathcal{O}(x) = \left\{\sigma^k(x), \quad k \in \mathbb{N}\right\} = \left\{x, \sigma(x), \ldots, \sigma^{p-1}(x)\right\}$$

Considérons

$$l = \inf\left\{k \in \mathbb{N}^* \mid \sigma^k(x) \in \left\{x, \sigma(x), \ldots, \sigma^{k-1}(x)\right\}\right\}$$

On a évidemment $l \leqslant p$ puisque $\sigma^p(x) = x$, et le caractère minimal de l montre que

$$x, \sigma(x), \ldots, \sigma^{l-1}(x) \text{ sont distincts deux à deux et } \sigma^l(x) = x$$

puisque, si $\sigma^l(x)$ était égal à $\sigma^k(x)$ avec $1 \leqslant k \leqslant l-1$, on obtiendrait $\sigma^{l-k}(x) = x$ (par injectivité de σ), en contradiction avec la définition de l. On en déduit que, si $l \geqslant 2\sigma$ opère sur $\{x, \sigma(x), \ldots, \sigma^{l-1}(x)\}$ comme le cycle $(x, \sigma(x), \ldots, \sigma^{l-1}(x))$ (avec $l = 1$ ssi $\mathcal{O}(x)$ est de cardinal 1, ce qui équivaut évidemment à $\sigma(x) = x$). Comme nous avons supposé que $\sigma \neq id$, il existe au moins une orbite de cardinal $\geqslant 2$. S'il y en a m, notons les $(\mathcal{O}(x_i))_{1 \leqslant i \leqslant m}$ de cardinaux respectifs $(l_i)_{1 \leqslant i \leqslant m}$. Ces orbites sont deux à deux disjointes et on a alors

$$\sigma = \prod_{i=1}^{m} \left(x_i, \sigma(x_i), \ldots, \sigma^{l_i - 1}(x_i)\right)$$

(le produit correspond ici à la composition des applications), ces différents cycles commutant deux à deux, puisqu'opérant sur des ensembles disjoints. Pour vérifier cette égalité, il suffit de montrer que tout élément x de $\{1, \ldots, n\}$ a même image par les deux membres. C'est évident si $x \notin \bigcup_{i=1}^{m} \mathcal{O}(x_i)$ (on a alors $\sigma(x) = x$) et c'est clair si $x = \sigma^k(x_i) \in \mathcal{O}(x_i)$. On a ainsi obtenu une décomposition de σ en produit de cycles "disjoints". Si réciproquement

$$\sigma = c_1 \circ \cdots \circ c_q$$

est une décomposition en cycles disjoints, avec $c_1 = (a_1, a_2, \ldots, a_{n_1})$ cycle de longueur n_1, il est clair que $\mathcal{O}(a_1) = \{a_1, a_2, \ldots, a_{n_1}\}$, et que le cycle c_1 apparaît dans la décomposition précédemment obtenue de σ. L'unicité en découle facilement. ■

EXEMPLE 17-2.24 En étudiant les orbites des différents éléments de $\{1, \ldots, 9\}$, montrer que, dans S_9

$$\sigma = \begin{pmatrix} 1,2,3,4,5,6,7,8,9 \\ 3,9,7,8,5,1,6,4,2 \end{pmatrix} = (1,3,7,6) \circ (2,9) \circ (4,8)$$

EXERCICE 17-2.25 Avec les notations de la démonstration qui précède, montrer que

$$p = [\sigma] = \mathrm{PPCM}(l_1, \ldots, l_m)$$

17-2.4 Quelques exemples d'applications

17-2.4.1 Classes de conjugaison

Lorsqu'on fait opérer un groupe $(\mathbb{G},.)$ sur lui-même par automorphisme intérieur (*cf* section **17-2.2.2**), deux éléments $x, y \in \mathbb{G}$ sont dits conjugués s'ils appartiennent à la même orbite, c'est-à-dire s'il existe un automorphisme intérieur transformant l'un en l'autre. L'orbite d'un point est appelée **classe de conjugaison** :

$$\mathcal{O}(x) = \{axa^{-1}, \quad a \in \mathbb{G}\}$$

De même, on dira que deux sous-groupes $\mathbb{G}_1$ et $\mathbb{G}_2$ de $\mathbb{G}$ sont conjugués si et seulement si

$$\exists a \in \mathbb{G} \quad \mathbb{G}_2 = a.\mathbb{G}_1.a^{-1}$$

On pourra ainsi parler de la classe de conjugaison d'un sous-groupe de $\mathbb{G}$.

EXERCICE 17-2.26 Montrer que la classe de conjugaison du noyau d'un morphisme de groupe ne contient qu'un élément.

EXERCICE 17-2.27 Soit $c = (a_1, \dots, a_k) \in S_n$ un cycle de longueur k. Montrer que, pour $\sigma \in S_n$, $\sigma \circ c \circ \sigma^{-1}$ est un cycle de longueur k. Décrire la classe de conjugaison de c.

EXERCICE 17-2.28 Soient σ et $\sigma' \in S_n$. Donner une condition nécessaire et suffisante sur les décompositions de σ et σ' en produit de cycles disjoints pour que σ et σ' soient conjuguées.

17-2.4.2 Stabilisateur d'un point

Soit $(\mathbb{G},.)$ un groupe opérant sur un ensemble $\mathbb{X}$. Si $x \in \mathbb{X}$, on appelle stabilisateur de x l'ensemble

$$\mathbb{G}_x = \{\sigma \in \mathbb{G} \mid \sigma.x = x\}$$

EXERCICE 17-2.29 Montrer que $\mathbb{G}_x$ est un sous-groupe de $\mathbb{G}$.

EXERCICE 17-2.30 Montrer que, si $x, y \in \mathbb{G}$ avec $y \in \mathcal{O}(x)$, il existe $\sigma \in \mathbb{G}$ avec

$$\mathbb{G}_x = \sigma.\mathbb{G}_y.\sigma^{-1}$$

Deux éléments appartenant à la même orbite ont donc des stabilisateurs conjugués, et en particulier isomorphes.

EXEMPLE 17-2.31 On fait opérer le groupe $\mathcal{O}_n(\mathbb{R})$ de manière naturelle sur l'ensemble des sous-espaces vectoriels de $\mathbb{R}^n$. Montrer que le stabilisateur d'un s.e.v est isomorphe à un des groupes $\mathcal{O}_p(\mathbb{R}) \times \mathcal{O}_{n-p}(\mathbb{R})$ pour $0 \leqslant p \leqslant n$.

Si l'ensemble $\mathbb{X}$ est fini, ou si le groupe $\mathbb{G}$ l'est, il en est évidemment de même de l'orbite d'un élément quelconque de $\mathbb{X}$. Dans le cas général, la connaissance du stabilisateur de $x \in \mathbb{X}$ donne une information sur le cardinal de $\mathcal{O}(x)$:

EXERCICE 17-2.32 Montrer que, pour σ et $\sigma' \in \mathbb{G}$ et $x \in \mathbf{X}$

$$\sigma.x = \sigma'.x \Leftrightarrow \sigma^{-1}\sigma' \in \mathbb{G}_x$$

En déduire que $\mathcal{O}(x)$ est fini si et seulement si $\mathbb{G}_x$ est d'indice fini (*cf* remarque **17-2**.18) dans $\mathbb{G}$, et qu'on a alors

$$\operatorname{card} \mathcal{O}(x) = [\mathbb{G} : \mathbb{G}_x]$$

ce qui s'écrira bien sûr

$$\operatorname{card} \mathcal{O}(x) = \frac{[\mathbb{G}]}{[\mathbb{G}_x]} \text{ si } \mathbb{G} \text{ est fini}$$

Une conséquence importante de ce résultat est l'**équation aux classes**, lorsque $\mathbb{X}$ est un ensemble fini. Comme les différentes orbites des éléments de $\mathbb{X}$ réalisent une partition de cet ensemble, si

$$\mathbb{X} = \bigcup_{i=1}^{p} \mathcal{O}(x_i)$$

est cette partition, nous aurons

$$\operatorname{card} \mathbb{X} = \sum_{i=1}^{p} \operatorname{card} \mathcal{O}(x_i) = \sum_{i=1}^{p} [\mathbb{G} : \mathbb{G}_{x_i}]$$

La section qui suit donne un exemple remarquable d'utilisation de ce "principe des tiroirs" :

17-2.4.3 Exemple d'utilisation

Si $\mathbb{G}$ est un groupe, on appelle **centre** de $\mathbb{G}$ et on note $Z(\mathbb{G})$ l'ensemble des éléments de $\mathbb{G}$ commutant avec tous les autres :

$$Z(\mathbb{G}) = \{x \in \mathbb{G} \mid \forall a \in \mathbb{G} \quad xa = ax\}$$

On vérifie aisément qu'il s'agit d'un sous-groupe[10] de $\mathbb{G}$. Dire que $\mathbb{G}$ est commutatif revient à $Z(\mathbb{G}) = \mathbb{G}$. Pour parler de manière imprécise, plus $Z(\mathbb{G})$ est "petit", moins le groupe est commutatif. On conçoit bien que les automorphismes intérieurs sont un outil d'étude de cette non-commutativité.

Si p est un nombre premier, un groupe fini $\mathbb{G}$ est un p-groupe ssi son cardinal est une puissance de p :

$$\exists m \in \mathbb{N}^* \quad [\mathbb{G}] = p^m$$

Le but de cette section est de prouver que le centre d'un p-groupe n'est jamais réduit à l'élément neutre. Pour cela, on fait opérer $\mathbb{G}$ sur lui-même par automorphismes intérieurs. Montrer que les orbites réduites à un point sont celles des éléments du centre. Montrer qu'une orbite non réduite à un point a pour cardinal une puissance de p (*cf* exercice **17-2**.32). En écrivant l'équation aux classes, montrer que $\operatorname{card} Z(\mathbb{G})$ est divisible par p.

[10] C'est $\bigcap_{a \in \mathbb{G}} \mathbb{G}_a$ si l'on fait opérer $\mathbb{G}$ sur lui-même par automorphismes intérieurs.

EXERCICE 17-2.33 Soit $\mathbb{G}$ un groupe d'ordre p^2, avec p premier. Si $\mathbb{G}$ n'est pas commutatif, montrer que $[Z(\mathbb{G})] = p$. Si $x \in \mathbb{G} - Z(\mathbb{G})$, montrer que l'ensemble des éléments de $\mathbb{G}$ commutant avec x est un sous-groupe de $\mathbb{G}$ contenant $\{x\} \cap Z(\mathbb{G})$. Montrer que c'est $\mathbb{G}$, et en déduire que $\mathbb{G}$ est commutatif. Il existe par contre des groupes d'ordre p^3 non commutatifs. Avec $p = 2$ nous avons l'exemple **17-1.31**

EXERCICE 17-2.34 Soit $\mathbb{G}$ un groupe (commutatif) d'ordre p^2, avec p premier. Montrer que, si $\mathbb{G}$ n'est pas cyclique, tout élément $x \neq 1_{\mathbb{G}}$ de $\mathbb{G}$ engendre un sous-groupe d'ordre p. Soit alors $y \in \mathbb{G}-\operatorname{Gr}(x)$. Montrer que les éléments

$$(x^i y^j)_{\substack{0 \leqslant i < p \\ 0 \leqslant j < p}}$$

sont distincts deux à deux, et en déduire un isomorphisme entre $(\mathbb{G}, .)$ et $\left((\mathbb{Z}/p\mathbb{Z})^2, +\right)$. Il n'y a donc, à isomorphisme près, que deux groupes d'ordre p^2.

17-2.5 Rappels sur le groupe symétrique

Nous rappelons ici les démonstrations de résultats évoqués dans le chapitre sur les déterminants.

DÉFINITION 17-2.35 *Si $\sigma \in S_n$ et $i, j \in \{1, \dots, n\}$, on dit que σ présente une inversion pour le couple (i, j) si et seulement si*

$$i < j \text{ et } \sigma(i) > \sigma(j)$$

On appelle signature de la permutation σ l'entier $\varepsilon(\sigma) \in \{\pm 1\}$ défini par

$$\varepsilon(\sigma) = (-1)^{\text{nombre d'inversions présentées par } \sigma}$$

La permutation σ est dite paire si $\varepsilon(\sigma) = +1$, elle est impaire dans le cas contraire.

PROPOSITION 17-2.36 Si $\sigma \in S_n$, on a

$$\varepsilon(\sigma) = \prod_{\substack{(i,j) \\ i<j}} \frac{\sigma(j) - \sigma(i)}{j - i}$$

La signature est un morphisme de groupes de $(S_n, \circ)$ dans $(\{\pm 1\}, \times)$.

Démonstration : Il est évident que le produit envisagé a même signe que $\varepsilon(\sigma)$. Il est de plus de valeur absolue égale à 1, puisque

$$\left| \prod_{\substack{(i,j) \\ i<j}} \frac{\sigma(j) - \sigma(i)}{j - i} \right| = \frac{\prod_{i<j} |\sigma(j) - \sigma(i)|}{\prod_{i<j} |j - i|}$$

et que le numérateur est égal au dénominateur, puisqu'on y retrouve les mêmes facteurs dans un ordre différent.

Si σ et σ' sont deux éléments de S_n, on a donc

$$\varepsilon(\sigma\circ\sigma') = \prod_{\substack{(i,j)\\ i<j}} \frac{\sigma\sigma'(j)-\sigma\sigma'(i)}{j-i}$$

$$= \prod_{\substack{(i,j)\\ i<j}} \frac{\sigma\sigma'(j)-\sigma\sigma'(i)}{\sigma'(j)-\sigma'(i)} \times \prod_{\substack{(i,j)\\ i<j}} \frac{\sigma'(j)-\sigma'(i)}{j-i}$$

Le second terme du produit vaut $\varepsilon(\sigma')$. Comme

$$\frac{\sigma[\sigma'(j)]-\sigma[\sigma'(i)]}{\sigma'(j)-\sigma'(i)} = \frac{\sigma[\sigma'(i)]-\sigma[\sigma'(j)]}{\sigma'(i)-\sigma'(j)}$$

et que σ' est une permutation de $\{1,\ldots,n\}$, en réordonnant les indices, on obtient

$$\prod_{\substack{(i,j)\\ i<j}} \frac{\sigma\sigma'(j)-\sigma\sigma'(i)}{\sigma'(j)-\sigma'(i)} = \prod_{\substack{(i',j')\\ i'<j'}} \frac{\sigma(j')-\sigma(i')}{j'-i'} = \varepsilon(\sigma)$$

ce qui donne finalement $\varepsilon(\sigma\circ\sigma') = \varepsilon(\sigma)\,\varepsilon(\sigma')$. ■

PROPOSITION 17-2.37 Toute transposition est une permutation impaire.

Démonstration : On peut directement (exercice) compter le nombre d'inversions d'une transposition $\tau = (i,j)$. On peut aussi remarquer que $\tau_0 = (1,2)$ présente une unique inversion (évidemment pour le couple $(1,2)$...) et donc $\varepsilon(\tau_0) = -1$, puis que, pour $i \neq j \in \{1,\ldots,n\}$ et $\tau = (i,j)$, on peut trouver $\sigma \in S_n$ avec $\tau = \sigma\circ\tau_0\circ\sigma^{-1}$ (*cf* exercice **17-2**.27) et donc $\varepsilon(\tau) = \varepsilon(\sigma)\,\varepsilon(\tau_0)\,[\varepsilon(\sigma)]^{-1} = \varepsilon(\tau_0) = -1$. ■

Dans la définition de la signature donnée plus haut, l'ordre "naturel" sur $\{1,\ldots,n\}$ semble avoir son importance. La proposition précédente et celle qui suit montreront qu'il n'en est rien :

PROPOSITION 17-2.38 Les transpositions engendrent le groupe $(S_n,\circ)$: toute permutation différente de l'identité[11] peut s'écrire comme produit de transpositions, et le nombre de ces transpositions peut être pris inférieur à $n-1$. Il n'y a pas unicité de cette décomposition, et ces transpositions ne commutent pas entre elles en général.

Démonstration : On peut (exercice) raisonner par récurrence sur n en montrant que, pour une permutation $\sigma \in S_n$ vérifiant $\sigma(n) \neq n$, on peut trouver une transposition τ telle que $\tau\circ\sigma(n) = n$, et on peut alors travailler avec la restriction de $\tau\circ\sigma$ à $\{1,\ldots,n-1\}$ qui

[11] Si τ est une transposition quelconque, on a évidemment $id_{\{1,\ldots,n\}} = \tau\circ\tau$.

est évidemment dans S_{n-1}. On peut aussi utiliser le théorème **17-2.23** de décomposition en cycles disjoints et remarquer qu'un cycle de longueur k peut se décomposer en $k-1$ transpositions :

$$(a_1, a_2, \ldots, a_k) = (a_1, a_k) \circ (a_1, a_{k-1}) \circ \cdots \circ (a_1, a_2)$$

(considérer l'image d'un a_i par le produit de transpositions composant le second membre de l'égalité, en se souvenant qu'on compose de droite à gauche). Il n'y a pas unicité de la décomposition, puisqu'on vérifie aussi

$$(a_1, a_2, \ldots, a_k) = (a_1, a_2) \circ (a_2, a_3) \circ \cdots \circ (a_k, a_{k-1}) \quad \blacksquare$$

On en déduit une caractérisation de la signature ne faisant pas intervenir d'ordre sur l'ensemble $\{1, \ldots, n\}$:

COROLLAIRE 17-2.39 Dans toutes les écritures d'une permutation comme produit de transpositions, la parité du nombre de transpositions est la même. Une permutation est paire si et seulement si elle peut s'écrire comme produit d'un nombre pair de transpositions. Une permutation impaire est produit d'un nombre impair de transpositions.

Démonstration : Si $\sigma \in S_n$ est produit de k transpositions

$$\sigma = \tau_1 \circ \cdots \circ \tau_k$$

on a $\varepsilon(\sigma) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_k) = (-1)^k$, puisque toute transposition est impaire. On en déduit la parité de k en fonction de la signature de σ. $\blacksquare$

La décomposition d'un cycle donnée plus haut donne alors

COROLLAIRE 17-2.40 Un cycle de longueur k est pair si et seulement si k est impair.

EXERCICE 17-2.41 Montrer qu'il existe exactement deux morphismes de groupes de $(S_n, \circ)$ dans $(\mathbb{R}^*, \times)$

DÉFINITION 17-2.42 *On appelle groupe alterné $(A_n, \circ)$ le noyau du morphisme*

$$\varepsilon : (S_n, \circ) \to (\{\pm 1\}, \times)$$

C'est un sous-groupe de $(S_n, \circ)$, formé de toutes les permutations paires. Il est de cardinal $\frac{n!}{2}$.

Il y a en effet deux classes (à droite si l'on veut) modulo A_n dans le groupe S_n : si τ est une transposition, τS_n contient toutes les permutations impaires. On en déduit aisément le cardinal de A_n.

17-3 Anneau $(\mathbb{Z}/n\mathbb{Z},+,\times)$

17-3.1 Structure d'anneau de $\mathbb{Z}/n\mathbb{Z}$

Si n est un entier $\geqslant 2$, on vérifie aisément que la multiplication interne dans $\mathbb{Z}$ est compatible avec la relation de congruence modulo n : si x, y, x', y' sont des entiers

$$\exists\, k, k' \in \mathbb{Z} \quad y = x + kn \text{ et } y' = x' + k'n \Rightarrow yy' = xx' + (kx' + k'x + kk'n)\,n$$

On peut alors définir dans l'ensemble quotient $\mathbb{Z}/n\mathbb{Z}$ une multiplication interne par

$$\overline{a}.\overline{b} \underset{\text{déf}}{=} \overline{ab}$$

et, comme cela a été vu à la section **17-1.3.1**, l'application $p : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ définie par $p(x) = \overline{x}$ est un morphisme (surjectif) à la fois pour les lois additives et multiplicatives, qui transporte donc la structure d'anneau :

PROPOSITION 17-3.1 $(\mathbb{Z}/n\mathbb{Z},+,\times)$ **est un anneau commutatif et**

$$p : (\mathbb{Z},+,\times) \to (\mathbb{Z}/n\mathbb{Z},+,\times) \textbf{ définie par } x \mapsto p(x) = \overline{x}$$

est un morphisme surjectif d'anneaux.

La classe $\overline{1}$ est évidemment l'élément unité de cet anneau.

Les calculs multiplicatifs dans $\mathbb{Z}/n\mathbb{Z}$ se font comme dans le cas de l'addition : on fait des multiplications sur des représentants des classes et on conserve du résultat son reste dans la division euclidienne par n. Par exemple, la table de $(\mathbb{Z}/4\mathbb{Z},\times)$ est

$\times$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

De nombreux résultats de divisibilité par n trouvent un traduction simple et efficace en travaillant dans cet anneau :

EXEMPLE 17-3.2 Montrer que, si n est un entier $\geqslant 1$

$$21 \text{ divise } a_n = 2^{4^n} + 5$$

Cela revient évidemment à montrer que la classe de a_n dans $\mathbb{Z}/21\mathbb{Z}$ est la classe de 0. $\mathbb{Z}/21\mathbb{Z}$ contient 21 éléments, ce qui rend un peu pénibles les calculs modulo 21. Il est plus simple de remarquer que $21 = 3 \times 7$ étant la décomposition de 21 en facteurs premiers, on doit montrer que 3 et 7 divisent a_n. On montre donc que les classes modulo 3 et modulo 7 de a_n sont nulles.

Dans $\mathbb{Z}/3\mathbb{Z}$, on a $\overline{2}^2 = \overline{1}$, ce qui entraîne évidemment $\overline{2}^4 = \overline{1}$ et $\overline{2^{4^n}} = \overline{2}^{4^n} = \left(\overline{2}^4\right)^{4^{n-1}} = \overline{1}$ d'où l'on déduit $\overline{a_n} = \overline{1} + \overline{5} = \overline{0}$.

De même, dans $\mathbb{Z}/7\mathbb{Z}$, on vérifie aisément $\widetilde{2}^4 = \widetilde{2}$, donc par récurrence sur n, $\widetilde{2}^{4^n} = \widetilde{2}$, et $\widetilde{a_n} = \widetilde{2} + \widetilde{5} = \widetilde{0}$, ce qui prouve le résultat.

EXEMPLE 17-3.3 Donner les deux derniers chiffres de l'écriture décimale de 3^{1994}.

Il s'agit ici de trouver le représentant canonique (compris entre 0 et 99) de 3^{1994} dans $\mathbb{Z}/100\mathbb{Z}$. Nous verrons à la section **17-3.3.3** que, dans $\mathbb{Z}/100\mathbb{Z}$

$$\overline{3}^{40}=\overline{1}$$

(ceci peut dès à présent se vérifier "à la main" : les congruences suivantes étant évidemment modulo 100, on a $3^5 = 81 \times 3 = 243 \equiv 43$, $43^2 = 1849 \equiv 49$ et $49^2 = 2401 \equiv 1$, ce qui donne en fait $3^{20} = \left((3^5)^2\right)^2 \equiv 1$). On travaille donc avec le reste de la division euclidienne de 1994 par 20 (de préférence à 40) : $1994 = 20 \times 99 + 14$, ce qui donne dans $\mathbb{Z}/100\mathbb{Z}$

$$\overline{3}^{1994}=\overline{3}^{20\times 99}.\overline{3}^{14}=\overline{3}^{14}$$

et comme $3^7 = 3^5.9 \equiv 43 \times 9 = 387 \equiv 87$, on a

$$\overline{3}^{14}=\overline{87}^2=\overline{87^2}=\overline{7569}=\overline{69}$$

les deux derniers chiffres cherchés sont 6 et 9. L'intérêt des méthodes "modulaires" est évident sur cet exemple : tous les calculs sont menés sur des entiers de tailles "raisonnables".

17-3.2 Groupe des unités de $\mathbb{Z}/n\mathbb{Z}$

17-3.2.1 Eléments inversibles pour la multiplication

Nous déterminons dans cette section les classes modulo n qui sont inversibles pour la multiplication. Comme cela a été vu en toute généralité dans un anneau quelconque, elles formeront un groupe pour la multiplication, le groupe des unités de l'anneau $(\mathbb{Z}/n\mathbb{Z},+,\times)$.

THÉORÈME 17-3.4 Si k est un entier relatif, sa classe $\overline{k}$ modulo n est inversible pour la multiplication dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si k est premier avec n. Les unités de l'anneau $(\mathbb{Z}/n\mathbb{Z},+,\times)$ sont donc exactement les générateurs du groupe $(\mathbb{Z}/n\mathbb{Z},+)$ (cf. section 17-1.3.4).

Démonstration : La classe de k est inversible si et seulement si

$$\exists u \in \mathbb{Z} \quad \overline{k}.\overline{u}=\overline{1}$$

soit, comme $\overline{k}.\overline{u}=\overline{ku}$

$$\exists u \in \mathbb{Z} \quad \exists v \in \mathbb{Z} \quad ku = 1 + nv$$

ce qui revient exactement à une identité de Bézout entre k et n. ■

On retiendra que, pour déterminer l'inverse d'une classe modulo n, on est amené à rechercher une identité de Bézout entre n et un représentant de cette classe. On se rappellera que l'algorithme d'Euclide est un moyen d'obtenir une telle identité : lorsque $k \wedge n = 1$, lorsqu'on effectue l'algorithme d'Euclide, les restes intermédiaires sont tous dans $k\mathbb{Z} + n\mathbb{Z}$, et le dernier reste non nul vaut 1.

REMARQUE 17-3.5 Si, à l'inverse, k n'est pas premier avec n, sa classe modulo n est **diviseur de zéro** dans $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ (et ne peut donc évidemment pas être inversible) : si

$$k \wedge n = d > 1$$

on peut écrire

$$\exists\, k', n' \in \mathbb{Z} \quad k' \wedge n' = 1 \text{ et } \begin{cases} n = d.n' \\ k = d.k' \end{cases}$$

Comme $0 < n' < n$, on a $\overline{n'} \neq \overline{0}$ et cependant

$$\overline{k}.\overline{n'} = \overline{d.k'.n'} = \overline{n.k'} = \overline{0}$$

En fait, la remarque qui précède est un cas particulier d'un résultat général :

EXERCICE 17-3.6 Si $(\mathbb{A}, +, \times)$ est un anneau fini, montrer que tout élément non inversible de $\mathbb{A}$ est diviseur de zéro dans $\mathbb{A}$.

17-3.2.2 Corps $\mathbb{Z}/p\mathbb{Z}$, p premier

THÉORÈME 17-3.7 Si $n \geqslant 2$ est un entier naturel, l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps si et seulement si n est un nombre premier.

Démonstration : Si n est premier, tout entier naturel k tel que $1 \leqslant k \leqslant n-1$ est premier avec n. Tout élément non nul de $\mathbb{Z}/n\mathbb{Z}$ est donc inversible pour la multiplication, et $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps.

Réciproquement, si n n'est pas premier, tout diviseur strict de n est diviseur de zéro dans $(\mathbb{Z}/n\mathbb{Z}, +, \times)$, qui ne peut par conséquent pas être un corps. ■

Si n est un entier $\geqslant 2$, il y a donc équivalence entre les propriétés

- n est premier.
- $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est intègre.
- $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps.

Cette structure de corps fini $(\mathbb{Z}/p\mathbb{Z}, +, \times)$, p premier, est d'une importance capitale dans de nombreux domaines en mathématique et informatique. Donnons un exemple d'utilisation :

EXEMPLE 17-3.8 (Théorème de Wilson) : si p est un entier naturel $\geqslant 2$, montrer que p est premier si et seulement

$$p \text{ divise } (p-1)! + 1$$

Si p divise $(p-1)! + 1$, il est clair qu'un diviseur d de p vérifiant $1 \leqslant d < p$ divise $(p-1)!$, et ne peut diviser $(p-1)! + 1$ que si $d = 1$, ce qui entraîne que p est

premier. Réciproquement, si p est premier $\geqslant 3$ (le cas $p=2$ est évident), on doit montrer que, dans $\mathbb{Z}/p\mathbb{Z}$,

$$\overline{(p-1)!} = -\overline{1}$$

Or on a, par définition de la multiplication dans $\mathbb{Z}/p\mathbb{Z}$,

$$\overline{(p-1)!} = \prod_{i=1}^{p-1} \overline{i}$$

qui est le produit de tous les éléments non nuls de $\mathbb{Z}/p\mathbb{Z}$. Il y a dans $(\mathbb{Z}/p\mathbb{Z})^* = \mathbb{Z}/p\mathbb{Z} - \{\overline{0}\}$ deux éléments égaux à leurs inverses. Ce sont les solutions de l'équation

$$x^2 = \overline{1}$$

dans $\mathbb{Z}/p\mathbb{Z}$. Comme $x^2 - \overline{1} = (x-\overline{1})(x+\overline{1})$, et que $\mathbb{Z}/p\mathbb{Z}$ est intègre, ces deux éléments sont $\overline{1}$ et $-\overline{1}$. En regroupant dans le produit des éléments non nuls de $\mathbb{Z}/p\mathbb{Z}$ chaque élément avec son inverse, ce qui laisse $\overline{1}$ et $-\overline{1}$ seuls, on obtient

$$\prod_{i=1}^{p-1} \overline{i} = -\overline{1}$$

et prouve le théorème de Wilson (critère de primalité d'un entier qui n'a pas d'intérêt pratique puisqu'il fait intervenir $(p-1)!$, entier de taille considérable dès que p n'est pas très petit).

17-3.3 Applications

17-3.3.1 Petit théorème de Fermat

Il s'agit d'un résultat qui donne une condition nécessaire pour qu'un nombre p soit premier. Il peut donc être utilisé pour prouver la non-primalité de certains entiers.

THÉORÈME 17-3.9 Si p est un nombre premier et a est un entier premier avec p, on a

$$a^{p-1} \equiv 1 \mod p$$

Démonstration : Nous donnerons deux démonstrations de ce résultat. La première n'est qu'une application immédiate du théorème de Lagrange : si p est premier, le groupe multiplicatif $U(\mathbb{Z}/p\mathbb{Z}) = \mathbb{Z}/p\mathbb{Z} - \{\overline{0}\}$ des unités de $\mathbb{Z}/p\mathbb{Z}$ contient $p-1$ éléments. Si $a \wedge p = 1$, on a $\overline{a} \neq \overline{0}$, et le théorème de Lagrange dans $(U(\mathbb{Z}/p\mathbb{Z}), \times)$ donne

$$\overline{a}^{p-1} = \overline{1}$$

ce qui prouve le théorème de Fermat.

La seconde démonstration est basée sur le fait que, si p est premier, pour k entier vérifiant $1 \leqslant k \leqslant p-1$

$$p \text{ divise } \mathrm{C}_p^k$$

ce qui est conséquence de l'égalité

$$k.\mathrm{C}_p^k = p.\mathrm{C}_{p-1}^{k-1}$$

qui montre que p divise $k.\mathrm{C}_p^k$, et divise donc C_p^k (théorème de Gauss), puisque $k \wedge p = 1$. La formule du binôme montre alors aisément que

$$\forall x, y \in \mathbb{Z} \quad (x+y)^p \equiv x^p + y^p \mod p$$

ce qui permet de prouver (par récurrence sur n) que

$$\forall n \in \mathbb{N} \quad n^p \equiv n \mod p$$

ce qui donne alors

$$\overline{a}^p = \overline{a}$$

Si de plus $a \wedge p = 1$, alors $\overline{a}$ est inversible dans $\mathbb{Z}/p\mathbb{Z}$, et on obtient, après simplification

$$\overline{a}^{p-1} = \overline{1}$$

ce qui prouve le résultat. ■

REMARQUE 17-3.10 Si p est un nombre premier, $\mathbb{K} = \mathbb{Z}/p\mathbb{Z}$ est un corps, et on peut considérer l'anneau $\mathbb{K}[X]$ des polynômes formels à coefficients dans $\mathbb{K}$. Il y a dans cet anneau une division euclidienne, et la théorie de la divisibilité se développe dans cet anneau comme lorsque $\mathbb{K}$ est un sous-corps de $\mathbb{C}$. L'application

$$\phi : P \mapsto \widetilde{P}$$

qui, à un polynôme formel P fait correspondre la fonction polynôme qui lui est naturellement associée, est un morphisme d'anneau de $\mathbb{K}[X]$ dans $\mathbb{K}^{\mathbb{K}}$. Comme la division euclidienne montre que l'on a

$$\forall x \in \mathbb{K} \quad \widetilde{P}(x) = \overline{0} \Leftrightarrow X - x \text{ divise } P$$

le théorème de Gauss entraîne toujours que, si $a_1, \ldots, a_k$ sont k éléments de $\mathbb{K}$ distincts deux à deux,

$$\forall i \quad \widetilde{P}(a_i) = \overline{0} \Leftrightarrow \prod_{i=1}^{k} (X - a_i) \text{ divise } P$$

En particulier, un polynôme non formellement nul de degré m ne peut posséder plus de m racines dans $\mathbb{K}$. Ceci entraîne aussi que

$$\forall P, Q \in \mathbb{K}[X] \quad \widetilde{P} = \widetilde{Q} \Leftrightarrow \prod_{a \in \mathbb{K}} (X - a) \text{ divise } P - Q$$

le noyau du morphisme ϕ étant l'idéal de $\mathbb{K}[X]$ engendré par $P_0 = \prod_{a \in \mathbb{K}} (X - a)$.

Le corps $\mathbb{K}$ étant fini, l'égalité fonctionnelle des polynômes de $\mathbb{K}[X]$ n'entraîne plus l'égalité formelle.

La démonstration précédente du théorème de Fermat donne l'exemple d'un polynôme $X^p - X$ non formellement nul, mais fonctionnellement nul.

EXERCICE 17-3.11 Montrer que, dans $\mathbf{K}[X]$, avec $\mathbf{K} = \mathbf{Z}/p\mathbf{Z}$ (p premier), on a

$$\prod_{a \in \mathbf{K}} (X - a) = X^p - X$$

En déduire une autre démonstration du théorème de Wilson.

EXERCICE 17-3.12 Avec les notations qui précèdent, montrer que ϕ est une surjection : toute fonction $\mathbf{K} \to \mathbf{K}$ est polynomiale !

EXERCICE 17-3.13 On considère l'anneau $(\mathbf{Z}[X], +, \times)$ (attention : il n'est pas principal, $\mathbf{Z}$ n'est pas un corps). Montrer que, pour p premier, l'application

$$(\mathbf{Z}[X], +, \times) \to (\mathbf{Z}/p\mathbf{Z}[X], +, \times)$$
$$P = \sum_{i=0}^{m} a_i X^i \mapsto \overline{P} = \sum_{i=0}^{m} \overline{a_i} X^i$$

est un morphisme d'anneaux. Quel est son noyau ?

Ce morphisme peut être utilisé pour démontrer de manière élégante des résultats dans l'anneau $(\mathbf{Z}[X], +, \times)$:

EXERCICE 17-3.14 Un polynôme de $\mathbf{Z}[X]$ est dit primitif si ses coefficients sont premiers entre eux dans leur ensemble. Montrer que le produit de deux polynômes primitifs est primitif. (Indication : supposer que ce n'est pas vrai, et considérer un nombre premier p divisant tous les coefficients du polynôme produit).

EXERCICE 17-3.15 (critère d'Eisenstein) : Soit

$$P = X^m + \sum_{k=1}^{m} a_k X^{m-k}$$

un polynôme de $\mathbf{Z}[X]$ tel qu'il existe un nombre premier p vérifiant

$$\forall k = 1, \ldots, m \quad p \text{ divise } a_k \text{ et } p^2 \text{ ne divise pas } a_m$$

Montrer que P est irréductible dans $\mathbf{Z}[X]$ (c'est-à-dire que P ne possède pas de diviseur non constant ou de degré différent de celui de P).

17-3.3.2 Théorème chinois

Considérons deux entiers naturels a et $b \geqslant 2$, premiers entre eux. Si $x \in \mathbb{Z}$, notons $\overline{x}$ sa classe modulo a et $\widetilde{x}$ sa classe modulo b. L'application "naturelle"

$$\begin{aligned} \varphi : (\mathbb{Z},+,\times) &\to (\mathbb{Z}/a\mathbb{Z}\times\mathbb{Z}/b\mathbb{Z},+,\times) \\ x &\mapsto \varphi(x) = (\overline{x},\widetilde{x}) \end{aligned}$$

est un morphisme d'anneaux. Son noyau est un idéal de $\mathbb{Z}$. C'est

$$\ker\varphi = \left\{x\in\mathbb{Z} \mid (\overline{x},\widetilde{x}) = \left(\overline{0},\widetilde{0}\right)\right\}$$

et donc

$$x \in \ker\varphi \Leftrightarrow a \text{ et } b \text{ divisent } x \Leftrightarrow ab \text{ divise } x$$

puisque a et b sont supposés premiers entre eux. Le noyau de φ est donc exactement $ab.\mathbb{Z}$, et deux entiers x et x' ont même image par φ si et seulement si ils sont congrus modulo ab. Si l'on note $\widehat{x}$ la classe de x modulo ab, l'application

$$\begin{aligned} \widehat{\varphi} : (\mathbb{Z}/ab\mathbb{Z},+,\times) &\to (\mathbb{Z}/a\mathbb{Z}\times\mathbb{Z}/b\mathbb{Z},+,\times) \\ \widehat{x} &\mapsto \widehat{\varphi}(\widehat{x}) = (\overline{x},\widetilde{x}) \end{aligned}$$

est donc parfaitement définie, et est clairement un morphisme injectif d'anneaux (car son noyau est réduit à $\left\{\widehat{0}\right\}$). Comme l'ensemble de départ et l'ensemble d'arrivée ont même cardinal fini ab, $\widehat{\varphi}$ est un isomorphisme d'anneaux, ce qui prouve en particulier la surjectivité de φ, qui n'était pas évidente *a priori*[12]. Le raisonnement précédent se généralise évidemment au cas de n entiers $a_i \geqslant 2$, premiers entre eux deux à deux.

THÉORÈME 17-3.16 Si $(a_i)_{1\leqslant i\leqslant n}$ sont n entiers $\geqslant 2$ premiers entre eux deux à deux, l'application

$$\varphi : \left(\mathbb{Z}/\left(\prod_{i=1}^{n} a_i\right)\mathbb{Z},+,\times\right) \longrightarrow \prod_{i=1}^{n}(\mathbb{Z}/a_i\mathbb{Z},+,\times)$$

définie par

$$\widehat{\varphi} : \widehat{x} \mapsto (\overline{x},\ldots,\widetilde{x})$$

est un isomorphisme d'anneaux.

Nous avons en fait démontré le corollaire suivant, connu sous l'appellation de "théorème chinois" :

COROLLAIRE 17-3.17 Si $(a_i)_{1\leqslant i\leqslant n}$ sont n entiers $\geqslant 2$ premiers entre eux deux à deux, le système de congruences

$$\begin{cases} x \equiv y_1 \mod a_1 \\ \qquad\vdots \\ x \equiv y_n \mod a_n \end{cases}$$

[12] Voir aussi l'exercice **17-1.46**

possède des solutions dans $\mathbb{Z}$ quels que soient les entiers $(y_i)_{1 \leqslant i \leqslant n}$. L'ensemble de ces solutions est une classe de congruence modulo $\prod_{i=1}^{n} a_i$.

Démonstration : Il est d'abord clair que, si ce système possède une solution particulière x_0, les autres solutions sont les entiers x vérifiant

$$\begin{cases} x \equiv x_0 \mod a_1 \\ \quad \vdots \\ x \equiv x_0 \mod a_n \end{cases}$$

c'est-à-dire tels que

$$\forall i \quad a_i \text{ divise } (x - x_0)$$

soit encore, puisque les a_i sont premiers entre eux deux à deux, le produit $a_1 a_2 \cdots a_n$ divise $x - x_0$. L'ensemble des solutions est donc $\widehat{x_0}$, classe de congruence de x_0 modulo $a_1 a_2 \cdots a_n$. L'existence (et l'unicité d'ailleurs) de cette classe est assurée par le théorème qui précède. On a simplement

$$\widehat{x_0} = \widehat{\varphi}^{-1}(\overline{y_1}, \ldots, \widetilde{y_n}) \quad ■$$

Comment trouver explicitement une solution de ce système ? Ici encore, c'est l'algorithme d'Euclide qui peut être utilisé : dans le cas où $n = 2$, un entier x est solution de

$$\begin{cases} x \equiv y_1 \mod a_1 \\ x \equiv y_2 \mod a_2 \end{cases}$$

si et seulement s'il peut s'écrire

$$x = y_1 + k_1 a_1 = y_2 + k_2 a_2$$

avec k_1 et $k_2 \in \mathbb{Z}$. Il s'agit donc de trouver les couples (k_1, k_2) d'entiers vérifiant

$$k_1 a_1 - k_2 a_2 = y_2 - y_1$$

De tels entiers existent puisque

$$a_1.\mathbb{Z} + a_2.\mathbb{Z} = \mathbb{Z}$$

les entiers a_1 et a_2 étant premiers entre eux. L'algorithme d'Euclide fournissant un couple (u, v) avec $ua_1 - va_2 = 1$, on pourra prendre $k_1 = u.(y_2 - y_1)$ et $k_2 = v.(y_2 - y_1)$. On a ainsi trouvé un entier y_0 tel que

$$\begin{cases} x \equiv y_1 \mod a_1 \\ x \equiv y_2 \mod a_2 \end{cases} \Leftrightarrow x \equiv y_0 \mod a_1 a_2$$

Dans le cas où $n > 2$, on écrira alors

$$\begin{cases} x \equiv y_1 \mod a_1 \\ \quad \vdots \\ x \equiv y_n \mod a_n \end{cases} \Leftrightarrow \begin{cases} x \equiv y_0 \mod a_1 a_2 \\ x \equiv y_3 \mod a_3 \\ \quad \vdots \\ x \equiv y_n \mod a_n \end{cases}$$

ce qui amène à un système de $n - 1$ congruences, auquel on appliquera le même méthode, les entiers $a_1 a_2, a_3, \ldots, a_n$ étant encore premiers entre eux deux à deux.
■

17-3.3.3 Complément : indicatrice d'Euler

Si n est un entier $\geqslant 2$, on appelle indicateur d'Euler de n l'entier

$$\varphi(n) = \text{card}\,\{k \in \mathbb{N} \mid 1 \leqslant k \leqslant n \text{ et } k \wedge n = 1\}$$

C'est donc d'après la section **17-3.2.1** le cardinal de l'ensemble des éléments inversibles de l'anneau $\mathbb{Z}/n\,\mathbb{Z}$. C'est aussi le nombre de générateurs d'un groupe cyclique d'ordre n. L'application

$$\varphi : \mathbb{N} - \{0, 1\} \to \mathbb{N}^*$$

est appelée (fonction) indicatrice d'Euler.

Si p est un entier premier, on a $\varphi(p) = p - 1$. Plus généralement, si q est un entier $\geqslant 1$

$$\varphi(p^q) = p^q - p^{q-1} = p^q\left(1 - \frac{1}{p}\right)$$

puisque les entiers de $[1, p^q]$ non premiers avec p sont les multiples de p, qui sont au nombre de p^{q-1} : ce sont

$$1.p, 2.p, \ldots, p^{q-1}.p$$

Une conséquence du théorème **17-3**.16 est que **l'indicatrice d'Euler est multiplicative :**

THÉORÈME 17-3.18 Si a et b sont deux entiers $\geqslant 2$ premiers entre eux

$$\varphi(ab) = \varphi(a).\varphi(b)$$

Démonstration : Ceci pourrait se prouver par un raisonnement direct. Le théorème chinois donne une démonstration élégante. L'application

$$\begin{aligned} \psi : (\mathbb{Z}/ab\mathbb{Z}, +, \times) &\to (\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}, +, \times) \\ \widehat{x} &\mapsto \psi(\widehat{x}) = (\overline{x}, \widetilde{x}) \end{aligned}$$

étant un isomorphisme d'anneaux réalise une bijection entre les groupes d'unités

$$U(\mathbb{Z}/ab\mathbb{Z}) \text{ et } U(\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}) = U(\mathbb{Z}/a\mathbb{Z}) \times U(\mathbb{Z}/b\mathbb{Z})$$

(voir à quelle condition un élément du produit $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ est inversible). Le premier groupe contient $\varphi(ab)$ éléments. Le second a pour cardinal $\varphi(a).\varphi(b)$, ce qui démontre le résultat. ■

COROLLAIRE 17-3.19 Si $n = p_1^{q_1} p_2^{q_2} \cdots p_k^{q_k}$ est la décomposition en facteurs premiers de n, on a

$$\varphi(n) = \prod_{i=1}^{k} \varphi(p_i^{q_i}) = \prod_{i=1}^{k} p_i^{q_i}\left(1 - \frac{1}{p_i}\right) = n.\prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right)$$

Par exemple

$$\varphi(100) = 100.\left(1 - \frac{1}{2}\right).\left(1 - \frac{1}{5}\right) = 40$$

Le théorème qui suit (Euler) est la généralisation du petit théorème de Fermat lorsqu'on travaille avec un entier qui n'est pas premier :

THÉORÈME 17-3.20 Si $n \geqslant 2$ et $a \in \mathbb{Z}$ est premier avec n, on a

$$a^{\varphi(n)} \equiv 1 \mod n$$

Démonstration : On applique le théorème de Lagrange à $\overline{a}$ qui appartient au groupe multiplicatif $U(\mathbb{Z}/n\mathbb{Z})$, dont le cardinal est $\varphi(n)$. ■

Par exemple, l'ordre de $\overline{3}$ dans le groupe des unités de $\mathbb{Z}/100\mathbb{Z}$ est un diviseur de $\varphi(100) = 40$. Nous avons vu à l'exemple **17-3**.3 que cet ordre est égal à 20.

17-3.3.4 Caractéristique d'un anneau

Soit $(\mathbb{A}, +, \times)$ un anneau dont l'élément unité est noté $1_\mathbb{A}$ comme d'habitude.

DÉFINITION 17-3.21 *On dit que $\mathbb{A}$ est de caractéristique nulle si $1_\mathbb{A}$ est d'ordre infini dans le groupe additif $(\mathbb{A}, +)$. Dans le cas contraire, on appelle caractéristique de $\mathbb{A}$ l'ordre de $1_\mathbb{A}$ dans $(\mathbb{A}, +)$:*

$$\operatorname{carac} \mathbb{A} = \inf \{k \in \mathbb{N}^* \mid k.1_\mathbb{A} = 0_\mathbb{A}\}$$

Si $q = \operatorname{carac} \mathbb{A}$, on a alors

$$\forall k \in \mathbb{Z} \quad k.1_\mathbb{A} = 0_\mathbb{A} \Leftrightarrow q \text{ divise } k$$

La caractéristique de l'anneau est un annulateur non seulement pour l'élément unité, mais pour tout élément de $\mathbb{A}$, puisque les règles de calcul dans un anneau montrent que

$$\forall x \in \mathbb{A} \quad q.x = (q.1_\mathbb{A}).x = 0_\mathbb{A}.x = 0_\mathbb{A}$$

EXEMPLE 17-3.22 $\mathbb{Z}$ est de caractéristique nulle. La caractéristique de $\mathbb{Z}/m\mathbb{Z}$ est évidemment égale à m.

On peut dire que l'on retrouve, dans tout anneau, l'un des exemples qui précède :

PROPOSITION 17-3.23 Si $(\mathbb{A}, +, \times)$ est un anneau de caractéristique nulle, l'application

$$\mathbb{Z} \to \mathbb{A} \quad k \mapsto k.1_\mathbb{A}$$

est un morphisme injectif d'anneaux. $(\mathbb{A}, +, \times)$ contient donc le sous-anneau $\mathbb{Z}.1_\mathbb{A}$, isomorphe à $\mathbb{Z}$.

Si $(\mathbb{A}, +, \times)$ est de caractéristique $q > 0$, cette même application a pour noyau $q.\mathbb{Z}$. Deux entiers congrus modulo q ayant même image, l'application

$$\mathbb{Z}/q\mathbb{Z} \to \mathbb{A} \quad \overline{k} \mapsto k.1_{\mathbb{A}}$$

est parfaitement définie, et est un morphisme injectif. Le sous-anneau $\mathbb{Z}.1_{\mathbb{A}}$ est cette fois isomorphe à $\mathbb{Z}/q\mathbb{Z}$.

PROPOSITION 17-3.24 Si $(\mathbb{A}, +, \times)$ est un anneau intègre, sa caractéristique est nulle ou est un nombre premier.

Démonstration : Si $q = \text{carac}\,\mathbb{A} \neq 0$, $\mathbb{A}$ contient un sous-anneau isomorphe à $\mathbb{Z}/q\mathbb{Z}$. Pour que cet anneau soit intègre, il est nécessaire (et suffisant) que q soit premier. On aurait aussi pu dire que, si $q = q_1 q_2$ est une factorisation de q,

$$q.1_{\mathbb{A}} = (q_1.1_{\mathbb{A}}).(q_2.1_{\mathbb{A}}) = 0_{\mathbb{A}}$$

entraîne, par intégrité de $\mathbb{A}$, $q_1.1_{\mathbb{A}} = 0_{\mathbb{A}}$ ou $q_2.1_{\mathbb{A}} = 0_{\mathbb{A}}$, soit q divise q_1 ou q_2, ce qui prouve bien que q est premier. ■

COROLLAIRE 17-3.25 La caractéristique d'un corps est nulle ou est un nombre premier.

Un corps de caractéristique p premier contient donc le sous-corps $\mathbb{Z}.1_{\mathbb{A}}$, isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

EXERCICE 17-3.26 Si $\mathbb{K}$ est un corps de caractéristique nulle, montrer que $\mathbb{K}$ contient un sous-corps isomorphe à $\mathbb{Q}$.

EXERCICE 17-3.27 Si $\mathbb{K}$ est un corps (commutatif) fini, montrer que son cardinal est une puissance d'un nombre premier (utiliser la notion d'espace vectoriel de dimension finie).

17-4 Exercices

EXERCICE 17-4.1 Soit G un groupe commutatif tel que

$$(\exists n \in \mathbb{N} - \{0\})\ (\forall x \in G)\ x^n = 1.$$

1. On suppose $n = ab$ avec $a \wedge b = 1$. On pose $G_a = \{x^a \mid x \in G\}$. Montrer que G_a est un sous-groupe de G. On définit de même G_b. Montrer que

$$(\forall x \in G)\ (\exists!\,(u, v) \in G_a \times G_b) \quad x = uv$$

2. On suppose n impair. Montrer que $x \to x^2$ est un automorphisme de G. Déterminer l'automorphisme réciproque. Même question avec $x \to x^k$ lorsque $k \wedge n = 1$.

EXERCICE 17-4.2 Soit G un groupe commutatif d'ordre pq où p et q sont deux nombres premiers distincts. Montrer que G est monogène.

EXERCICE 17-4.3 Trouver tous les morphismes de $(\mathbb{Q},+)$ dans $(\mathbb{Z},+)$.

EXERCICE 17-4.4 Soit G un groupe fini, H un sous-groupe de G et p le plus petit diviseur premier de card G. On suppose que $\dfrac{\text{card}G}{\text{card}H} = p$. Montrer que, pour tout $x \in G$, $xHx^{-1} = H$. (On fera opérer G par translation sur les classes à gauche modulo H).

EXERCICE 17-4.5 Soit $p \geqslant 3$ premier et $\alpha \in \mathbb{N}$. Montrer que

$$(1+p)^{p^\alpha} \equiv 1 + p^{\alpha+1} \quad [p^{\alpha+2}]$$

EXERCICE 17-4.6 Soit p un nombre premier impair et q un diviseur premier de $2^p - 1$. Montrer que

$$q \equiv 1 \quad [2p]$$

EXERCICE 17-4.7 Soit $(S_n, \circ)$ le groupe symétrique et σ un cycle de longueur l.

1. Montrer que toute permutation conjuguée à σ est un cycle de longueur l.
2. En déduire un condition nécessaire et suffisante portant sur les décompositions en cycles disjoints pour que deux permutations soient conjuguées.

EXERCICE 17-4.8 $(S_n, \circ)$ est le groupe symétrique et $s \in S_n$ se décompose en produit cycles disjoints avec, pour $1 \leqslant i \leqslant n$, k_i cycles de longueur i. Montrer qu'il y a $\prod_{i=1}^{n} (k_i)!\, i^{k_i}$ permutations dans S_n qui commutent avec s.

EXERCICE 17-4.9 Montrer que le groupe symétrique $(S_n, \circ)$ est engendré par la transposition $(1,2)$ et le cycle $(1, 2, \ldots, n)$.

EXERCICE 17-4.10 Soit G un groupe fini et $\varphi : G \to G$ un endomorphisme. Montrer que :

$$\text{Ker}(\varphi) = \text{Ker}(\varphi^2) \iff \text{Im}(\varphi) = \text{Im}(\varphi^2)$$

EXERCICE 17-4.11 Soit G un groupe de cardinal n et p un diviseur premier de n. On pose

$$H = \{x \in G \mid x^p = 1\}$$

Montrer que p divise le cardinal de H. (On pourra commencer à traiter le cas $p = 2$ et considérer $\{(x, y) \in G^2 \mid xy = 1\}$).

EXERCICE 17-4.12 Trouver tous les entiers $n \in \mathbb{N}^*$ tels que $\mathbb{Z}/n\mathbb{Z}$ possède au moins un élément nilpotent non nul.

EXERCICE 17-4.13 Polynômes cyclotomiques : Pour tout entier $n \geqslant 1$, on appelle polynôme cyclotomique d'indice n le polynôme

$$F_n(X) = \prod_{\omega \in \Omega_n} (X - \omega)$$

où Ω_n est l'ensemble des racines primitives $n^{\text{ème}}$ de l'unité. On pose $F_1(X) = X - 1$.

1. Montrer que, si p est un entier premier, alors $F_p(X) = 1 + X + X^2 + \cdots + X^{p-1}$.

2. Montrer que, pour tout entier n, F_n est un polynôme (*a priori* à coefficients complexes) de degré $\varphi(n)$, φ étant l'indicatrice d'Euler. Calculer les 10 premiers polynômes cyclotomiques.

3. Montrer que, pour tout n, on a

$$X^n - 1 = \prod_{d \text{ divisant } n} F_d(X)$$

(on utilise tous les diviseurs positifs de n, 1 et n compris).

4. En déduire que tous les F_n sont des polynômes à coefficients entiers.

EXERCICE 17-4.14 Pour $n \in \mathbb{N}$ soit $a_n = 2^{2^n} + 1$. Montrer que

$$n \neq m \Longrightarrow a_n \wedge a_m = 1$$

EXERCICE 17-4.15 Trouver les $y \in \mathbb{N}^*$ tels que $y^y \equiv 1$ [7].

Chapitre 18

Calcul différentiel en dimension finie

18-1 Différentiabilité. Fonctions de classe $\mathcal{C}^1$

Dans tout ce chapitre, $\mathbb{E}_n$ représente un $\mathbb{R}$-ev de dimension n muni d'une norme quelconque, et $(\mathbb{F}, \| \ \|)$ est un $\mathbb{R}$-ev normé. Le plus souvent, $\mathbb{F}$ sera de dimension finie p et on le notera alors $\mathbf{F} = \mathbb{F}_p$. Dans ce cas, le choix de la norme sur $\mathbb{F}$ n'aura pas à être précisé en général. Les applications considérées seront définies sur un ouvert U de $\mathbb{E}_n$, à valeurs dans $\mathbb{F}$.

18-1.1 Dérivée selon un vecteur. Dérivées partielles dans une base

18-1.1.1 Dérivée selon un vecteur

L'idée est ici d'étudier une fonction f au voisinage d'un point $a \in U$, dans une direction particulière définie par un vecteur (non nul) $v \in \mathbb{E}_n$, en considérant l'application

$$t \mapsto f(a + tv)$$

définie sur un voisinage de 0 dans $\mathbb{R}$.

DÉFINITION 18-1.1 *Soit $f : U \to \mathbb{F}$ et $a \in U$. Si $v \in \mathbb{E}_n$, on dit que la fonction f est dérivable en a selon le vecteur v si et seulement si la fonction $t \mapsto f(a+tv)$ est dérivable en 0. Le vecteur de $\mathbb{F}$ défini par*

$$D_v f(a) = \frac{d}{dt}(f(a+tv))\,|_{t=0} = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}$$

est alors appelé vecteur dérivé de f en a selon v (ou suivant v).

Cette définition n'a évidemment d'intérêt que si $v \neq 0_{\mathbb{E}_n}$. Si $y = \lambda v$ est un vecteur colinéaire à v, l'existence de $D_v f(a)$ entraîne celle de $D_y f(a)$, avec

$$D_{\lambda v} f(a) = \lambda D_v f(a)$$

C'est pour cela qu'on utilise aussi "dérivée dans la direction du vecteur v" pour désigner $D_v f(a)$, même si cette terminologie est moins précise, puisque $D_v f(a)$ dépend quand même (de façon homogène) du vecteur directeur de $\operatorname{vect}(v)$.

Le fait de considérer f de manière "radiale" autour de a et d'obtenir une dérivabilité donne bien sûr des informations importantes, mais n'assure pas par exemple la continuité de f en a :

EXEMPLE 18-1.2 La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} 1 \text{ si } |y| \geqslant x^2 \text{ ou si } y = 0 \\ 0 \quad \text{sinon} \end{cases}$$

est discontinue à l'origine, mais présente en ce point une dérivée nulle selon tout vecteur.

18-1.1.2 Dérivées partielles dans une base

Lorsqu'on considère une base $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ de l'espace de départ, on identifie souvent un vecteur quelconque $x = \sum_{i=1}^{n} x_i e_i$ avec le n-uplet $(x_1, \ldots, x_n)$ de ses coordonnées dans $\mathcal{B}$. Considérer $f : U \to \mathbb{F}$ revient alors à envisager une fonction $\tilde{f} : U' \to \mathbb{F}$, où U' est un ouvert de $\mathbb{R}^n$ et $\tilde{f}$ est définie par

$$\tilde{f}(x_1, \ldots, x_n) = f\left(\sum_{i=1}^{n} x_i e_i\right)$$

On privilégie alors les directions des axes de coordonnées, ce qui amène à la définition suivante :

DÉFINITION 18-1.3 *Soit $f : U \to \mathbb{F}$ et $a \in U$, avec $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ base de l'espace de départ. Si $i \in \{1, \ldots, n\}$, on dit que la fonction f possède en a une dérivée partielle par rapport à la $i^{\text{ème}}$ coordonnée dans $\mathcal{B}$ si et seulement si elle est dérivable selon le vecteur e_i. Le vecteur dérivé $D_{e_i} f(a)$ est alors appelé dérivée partielle de f en a par rapport à la $i^{\text{ème}}$ coordonnée dans $\mathcal{B}$. S'il n'y a pas d'ambiguïté sur cette base, on le note simplement $D_i f(a)$:*

$$D_i f(a) = D_{e_i} f(a) = \frac{d}{dt}(f(a+te_i))\,|_{t=0} = \lim_{t\to 0} \frac{f(a+te_i) - f(a)}{t}$$

Autre notation : Lorsque l'on travaille avec la fonction $\widetilde{f}$ définie plus haut, et en notant $(a_1, \dots, a_n)$ les coordonnées de a, on a

$$D_i f(a) = \lim_{h \to 0} \frac{\widetilde{f}(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - \widetilde{f}(a_1, \dots, a_n)}{h}$$

C'est donc la dérivée en a_i de la $i^{\text{ème}}$ application partielle définie par $\widetilde{f}$ en a, qui est l'application (définie au voisinage de a_i)

$$t \mapsto \widetilde{f}(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$$

Dans ce cas (fonction définie sur un ouvert de $\mathbb{R}^n$), plutôt que de parler de dérivée partielle par rapport à la $i^{\text{ème}}$ coordonnée dans la base canonique, on parle de dérivée partielle par rapport à la $i^{\text{ème}}$ variable. Si on note $(x_1, \dots, x_n) = x$ le point "générique" de U', cette dérivée partielle sera notée

$$D_i \widetilde{f}(a_1, \dots, a_n) = \frac{\partial \widetilde{f}}{\partial x_i}(a_1, \dots, a_n)$$

Remarquons que, dans cette écriture, le symbole x_i ne représente rien de particulier. $\dfrac{\partial \widetilde{f}}{\partial x_i}$ signifie simplement que l'on évalue la dérivée de la $i^{\text{ème}}$ application partielle de $\widetilde{f}$. Et comme l'usage (en Physique notamment) veut que l'on ne change pas nécessairement le nom d'une fonction lorsqu'on effectue un changement de repère, nous noterons également ainsi la dérivée partielle de f

$$D_i f(a) = \frac{\partial f}{\partial x_i}(a)$$

Cet abus d'écriture est commode, mais un peu dangereux. Il faut toujours comprendre la portée des symboles utilisés. Si on représente par

$$\frac{\partial}{\partial x_1}[g(x_1, x_2)]$$

le résultat de l'opérateur de dérivation par rapport à la variable x_1 sur une expression formelle $g(x_1, x_2)$, on aura par exemple

$$\frac{\partial}{\partial x_1}\left[e^{x_1} \cos\left(x_1^2 x_2\right)\right] = e^{x_1}\left[\cos\left(x_1^2 x_2\right) - 2x_1 x_2 \sin\left(x_1^2 x_2\right)\right]$$

ce qui signifie la même chose que

$$\frac{\partial}{\partial y}\left[e^y \cos\left(y^2 z\right)\right] = e^y\left[\cos\left(y^2 z\right) - 2yz \sin\left(y^2 z\right)\right]$$

Si $(x, y) \mapsto f(x, y)$ est une fonction de deux variables définie sur $\mathbb{R}^2$ admettant en tout point des dérivées partielles par rapport à chacune des variables, on aura de même

$$\frac{\partial}{\partial x_1}[f(x_2, x_1)] = \frac{\partial f}{\partial y}(x_2, x_1)$$

si on a décidé de noter $\dfrac{\partial f}{\partial y}$ la fonction "dérivée partielle de f par rapport à la deuxième variable" (en travaillant dans la base canonique de $\mathbb{R}^2$). On aurait d'ailleurs aussi, par application d'une formule de dérivée d'une fonction composée[1]

$$\frac{\partial}{\partial x_1}\left[f\left(x_2, x_1^2 x_2\right)\right] = 2x_1x_2\frac{\partial f}{\partial y}\left(x_2, x_1^2x_2\right)$$

Les choses se corseraient un peu si nous utilisions à nouveau les lettres x et y. On aurait alors

$$\frac{\partial}{\partial x}\left[f\left(y, x^2 y\right)\right] = 2xy\frac{\partial f}{\partial y}\left(y, x^2y\right), \text{ expression différente de } \frac{\partial f}{\partial x}\left(y, x^2y\right)$$

On mesure la difficulté de l'utilisation des symboles $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$ et $\dfrac{\partial}{\partial x}$. Une écriture comme

$$\frac{\partial}{\partial x}\left[f\left(y, x^2 y\right)\right] = 2xyD_2f\left(y, x^2y\right)$$

est sans doute préférable.

Lorsque l'espace d'arrivée est de dimension finie p, les dérivées selon un vecteur et les dérivées partielles se calculent évidemment coordonnée par coordonnée :

PROPOSITION 18-1.4 Si $\mathcal{B}' = (u_i)_{1\leqslant i\leqslant p}$ est une base de $\mathbb{F}_p$, une fonction $f : U \to \mathbb{F}_p$ est caractérisée par ses fonctions composantes $(f_i)_{1\leqslant i\leqslant p}$, avec $f_i : U \to \mathbb{R}$ et

$$\forall x \in U \quad f(x) = \sum_{i=1}^{p} f_i(x)\, u_i$$

qu'on note en abrégé $f = \sum_{i=1}^{p} f_i u_i$, ou encore $f = (f_1, \ldots, f_p)$ si on identifie $\mathbb{F}_p$ avec $\mathbb{R}^p$ par le choix de $\mathcal{B}'$. Il est clair que, pour $a \in U$ et $x \in \mathbb{E}_n$, f possède une dérivée en a selon x si et seulement si ses composantes dans $\mathcal{B}'$ possèdent la même propriété, avec alors

$$D_x f(a) = \sum_{i=1}^{p} D_x f_i(a)\, u_i$$

De même, pour les dérivées partielles, on aura

$$D_j f(a) = \sum_{i=1}^{p} D_j f_i(a)\, u_i = \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^{p} \frac{\partial f_i}{\partial x_j}(a)\, u_i$$

---

[1] Pour les fonctions d'une variable réelle. Pour évaluer $\dfrac{\partial}{\partial x_2}\left[f\left(x_2, x_1^2 x_2\right)\right]$, il faudra faire des hypothèses supplémentaires sur f, par exemple supposer f de classe C^1, et utiliser un théorème de dérivée d'une composée plus élaboré (voir théorème **18-2.3.1**)

Pour terminer, notons que l'existence en un point $a \in U$ de dérivées partielles par rapport à toutes les coordonnées dans une base n'implique pas l'existence en a de dérivée selon tout vecteur de $\mathbb{E}_n$. Considérer par exemple l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(0,0) = 0 \text{ et pour } (x,y) \neq 0 \quad f(x,y) = \frac{xy}{x^2+y^2}$$

En travaillant avec la base canonique de $\mathbb{R}^2$, on obtient aisément

$$D_1 f(0,0) = D_2 f(0,0) = 0$$

mais la dérivée selon le vecteur $x = (1,1)$ n'existe pas, puisque

$$\frac{f(t,t) - f(0,0)}{t} = \frac{1}{2t} \text{ n'a pas de limite en } 0$$

18-1.1.3 Fonctions de classe C^1 sur un ouvert.

DÉFINITION 18-1.5 *Une fonction $f : U \to \mathbb{F}_p$ est dite de classe C^1 sur l'ouvert U si et seulement si :*

1. *f admet en tout point de U une dérivée selon tout vecteur de $\mathbb{E}_n$.*
2. *Pour tout $x \in \mathbb{E}_n$, l'application $U \to \mathbb{F}_p$ définie par $a \mapsto D_x f(a)$ est continue sur U.*

Si f et $g : U \to \mathbb{F}_p$ possèdent en un point $a \in U$ une dérivée selon un vecteur x, il en est clairement de même de toute combinaison linéaire de f et g et on a

$$D_x(f + \lambda g)(a) = D_x f(a) + \lambda D_x g(a)$$

Si f et g sont des fonctions numériques ($\mathbb{F}_p = \mathbb{R}$ ou $\mathbb{C}$), on aura de même

$$D_x(fg)(a) = f(a) D_x g(a) + g(a) D_x f(a)$$

On en déduit facilement les propositions :

PROPOSITION 18-1.6 L'ensemble $C^1(U, \mathbb{F}_p)$ est un sous-espace vectoriel de l'espace des applications de U dans $\mathbb{F}_p$.

PROPOSITION 18-1.7 Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, l'ensemble $C^1(U, \mathbb{K})$ est une sous-algèbre de $\mathbb{K}^U$.

En particulier, les fonctions polynomiales de $\mathbb{R}^n \to \mathbb{R}$ sont dans $C^1(\mathbb{R}^n, \mathbb{R})$ puisque pour une telle fonction f, il est facile de voir que $a \to D_x f(a)$ est aussi une application polynomiale. De même, une fonction fraction rationnelle est de classe C^1 sur son ouvert de définition.

18-1.2 Différentiabilité en un point

18-1.2.1 Définition

DÉFINITION 18-1.8 *$f : U \to \mathbb{F}_p$ est dite différentiable en un point $a \in U$ si et seulement s'il existe une application linéaire $l \in \mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)$ telle que l'on ait, pour h tendant vers $0_{\mathbb{E}_n}$*

$$f(a+h) = f(a) + l(h) + o(\|h\|) \quad (*)$$

ce qui peut aussi s'écrire

$$f(a+h) = f(a) + l(h) + \|h\| \varepsilon(h) \quad \text{avec} \quad \lim_{h \to 0_{\mathbb{E}_n}} \varepsilon(h) = 0_{\mathbb{F}_p}$$

On dit aussi que f possède un développement limité d'ordre 1 en a. Lorsque $\mathbb{E}_n = \mathbb{R}$, on sait que ceci équivaut à la dérivabilité de f en a, avec

$$\forall h \in \mathbb{R} \quad l(h) = hf'(a)$$

PROPOSITION 18-1.9 Si $f : U \to \mathbb{F}_p$ est différentiable en a, elle est continue en a.

Démonstration : Comme l est une application linéaire définie sur un espace normé de dimension finie[2], elle est continue. On a donc, avec les notations précédentes

$$\lim_{h \to 0_{\mathbb{E}_n}} l(h) + \|h\| \varepsilon(h) = 0_{\mathbb{F}_p}$$

ce qui prouve bien la continuité de f en a. ■

PROPOSITION 18-1.10 Si $f : U \to \mathbb{F}_p$ est différentiable en a et vérifie $(*)$, elle possède en a une dérivée selon tout vecteur $x \in \mathbb{E}_n$, et on a

$$\forall x \in \mathbb{E}_n \quad D_x f(a) = l(x)$$

Cette dérivée dépend donc linéairement[3] du vecteur x.

Démonstration : Il existe un réel $\alpha > 0$ tel que

$$\|h\| \leqslant \alpha \Rightarrow a + h \in U$$

On a alors, pour t réel

$$|t| \leqslant \frac{\alpha}{\|x\|} \Rightarrow f(a+tx) = f(a) + tl(x) + |t| \|x\| \varepsilon(tx)$$

qui donne aisément

$$\lim_{t \to 0} \frac{f(a+tx) - f(a)}{t} = l(x) \quad ■$$

[3]C'est pour cela que, lorsque l'espace $\mathbb{E}$ de départ est de dimension infinie, on dira que $f : U \to \mathbb{F}$ est différentiable en un point a ssi il existe une application linéaire **continue** $l \in \mathcal{L}_c(\mathbb{E}, \mathbb{F})$ telle que

$$f(a+h) \underset{h \to 0_{\mathbb{E}}}{=} f(a) + l(h) + o(\|h\|)$$

[3]L'exemple **18-1.2** montre que cette seule propriété n'entraîne pas la continuité et *a fortiori* la différentiabilité de f.

COROLLAIRE 18-1.11 (et DÉFINITION) Si f est différentiable en $a \in U$, l'application linéaire $l \in \mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)$ qui intervient dans le développement limité de f est unique. On l'appelle différentielle de f en a, et on la note[4] df_a. On a donc

$$f(a+h) = f(a) + df_a(h) + o(\|h\|) \text{ pour } h \to 0_{\mathbb{E}_n}$$

L'application

$$g : x \to f(a) + df_a(x-a)$$

est appelée application affine tangente à f en a, ce qui signifie simplement que cette application affine vérifie

$$f(x) - g(x) \underset{x \to +a}{=} o(\|x-a\|)$$

EXEMPLE 18-1.12 Il est évident qu'une application linéaire $f \in \mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)$ est différentiable en tout point de $\mathbb{E}_n$ et que, pour tout $a \in \mathbb{E}_n$, on a $df_a = f$. De même une application affine $\mathbb{E}_n \to \mathbb{F}_p$ possède en tout point une différentielle égale à sa partie linéaire.

18-1.2.2 Expression dans une base

Regardons d'abord ce qui se passe lorsqu'on travaille avec une base $\mathcal{B}' = (u_i)_{1 \leqslant i \leqslant p}$ de l'espace d'arrivée $\mathbb{F}_p$. On écrit alors

$$f = \sum_{i=1}^{p} f_i u_i$$

Il est clair que f possède un développement limité en a si et seulement si chacune des applications coordonnées en possède un. On a alors

PROPOSITION 18-1.13 f est différentiable en a si et seulement si les $(f_i)_{1 \leqslant i \leqslant p}$ le sont, et on a

$$\forall h \in \mathbb{E}_n \quad df_a(h) = \sum_{i=1}^{p} (df_i)_a(h)\, u_i$$

Fixons à présent une base $\mathcal{B} = (e_j)_{1 \leqslant j \leqslant n}$ de $\mathbb{E}_n$. La fonction f devient alors une fonction de n variables réelles, que nous noterons $(x_1, \ldots, x_n)$. On a alors

[4]On note également $df_a = f'(a)$, et on écrit

$$f(a+h) = f(a) + f'(a).h + o(\|h\|)$$

où $f'(a).h$ représente l'image du vecteur h par le morphisme $f'(a)$. Il y a alors abus d'écriture dans le cas $n = 1$, où l'on confond le vecteur dérivé $f'(a)$ avec le morphisme $\mathbb{R} \to \mathbb{F}_p$ défini par $h \to hf'(a)$.

THÉORÈME 18-1.14 Si f est différentiable en a, elle possède en ce point des dérivées partielles par rapport à chacune des coordonnées dans $\mathcal{B}$, et on a, pour $h = \sum_{j=1}^{n} h_j ej$ vecteur quelconque de $\mathbb{E}_n$

$$df_a(h) = \sum_{j=1}^{n} h_j D_j f(a) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a)$$

Démonstration : L'existence de dérivées partielles est assurée par la proposition **18-1**.10, puisque

$$D_j f(a) = D_{e_j} f(a) = df_a(e_j)$$

et, comme df_a est linéaire

$$df_a\left(\sum_{j=1}^{n} h_i e_i\right) = \sum_{j=1}^{n} h_i df_a(e_i) = \sum_{j=1}^{n} h_i D_j f(a) \quad \blacksquare$$

Rappelons encore une fois que, réciproquement, l'existence de dérivées partielles de f en a par rapport à chaque coordonnée dans $\mathcal{B}$ n'entraîne pas nécessairement la différentiabilité de f en a.

On note souvent l'expression de la différentielle utilisant les dérivées partielles dans $\mathcal{B}$ sous la forme

$$df_a = \sum_{j=1}^{n} D_j f(a)\, dx_j = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, dx_j$$

où les $(dx_j)_{1 \leqslant j \leqslant n}$ représentent les formes coordonnées dans la base $\mathcal{B}$. Nous verrons ultérieurement l'intérêt de cette notation, notamment lorsqu'on fait des changements de variables.

18-1.2.3 Matrice Jacobienne

Soit $f : U \to \mathbb{F}_p$ différentiable en a. Si on choisit pour les espaces $\mathbb{E}_n$ et $\mathbb{F}_p$ des bases respectives $\mathcal{B}$ et $\mathcal{B}'$ pour représenter $f : U \to \mathbb{F}_p$

$$f\left(\sum_{j=1}^{n} x_i e_i\right) = \sum_{i=1}^{p} f_i(x_1, \ldots, x_n)\, u_i$$

il est naturel d'exprimer le morphisme df_a par sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$. Ceci amène à la définition

DÉFINITION 18-1.15 *Lorsque f est différentiable en a, on appelle matrice Jacobienne de f en a par rapport aux bases $\mathcal{B}$ et $\mathcal{B}'$ la matrice de la différentielle df_a dans les bases $\mathcal{B}$ et $\mathcal{B}'$:*

$$J^{\mathcal{B},\mathcal{B}'}(f)(a) = M(df_a, \mathcal{B}, \mathcal{B}') \in \mathcal{M}_{p,n}(\mathbb{R})$$

Lorsqu'il n'y a pas d'ambiguïté sur les bases, on notera simplement $J(f)(a)$ (ce qui est souvent le cas pour f définie d'un ouvert de $\mathbb{R}^n$ vers $\mathbb{R}^p$, où l'on travaille en général avec les bases canoniques). Avec les notations qui précèdent cette définition, on aura

$$J(f)(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_j}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial x_1}(a) & \cdots & \frac{\partial f_i}{\partial x_j}(a) & \cdots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \frac{\partial f_p}{\partial x_j}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

puisque la $j^{\text{ème}}$ colonne représente les coordonnées de $df_a(e_j) = D_j f(a)$ dans la base $\mathcal{B}'$.

Cette matrice représentant un morphisme d'espaces vectoriels, elle se transforme de manière évidente lors d'un changement de bases :

PROPOSITION 18-1.16 Si $\mathcal{B}_1$ et $\mathcal{B}'_1$ sont d'autres bases de $\mathbb{E}_n$ et $\mathbb{F}_p$, caractérisées par les matrices de passage $P \in \mathcal{GL}_n(\mathbb{R})$ (de $\mathcal{B}$ à $\mathcal{B}_1$) et $Q \in \mathcal{GL}_p(\mathbb{R})$ (de $\mathcal{B}'$ à $\mathcal{B}'_1$), pour f différentiable en a, la matrice Jacobienne par rapport à $\mathcal{B}_1$ et $\mathcal{B}'_1$ sera donnée par

$$J^1(f)(a) = Q^{-1} J(f)(a) P$$

Lorsque $\mathbb{E}_n = \mathbb{F}_p$, la différentielle df_a est un endomorphisme de $\mathbb{E}_n$. Son déterminant (qui peut être évalué dans une base quelconque de $\mathbb{E}_n$) est alors appelé (déterminant) Jacobien de f en a. Si

$$f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{i=1}^{n} f_i(x_1, \ldots, x_n) e_i$$

on aura

$$\det J(f)(a) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{vmatrix}$$

Lorsque l'espace $\mathbb{E}_n$ est euclidien, on sait que la valeur absolue de ce déterminant peut être interprétée comme un facteur de multiplication des volumes puisque, pour $y_1, \ldots, y_n$ indépendants dans $\mathbb{E}_n$,

$$\det J(f)(a) = \frac{[df_a(y_1), \ldots, df_a(y_n)]}{[y_1, \ldots, y_n]}$$

est un quotient de deux produit mixtes, donc est rapport du volume (algébrique) du parallélépipède construit sur les vecteurs $df_a(y_i)$ au volume de celui construit sur les y_i. On aura alors clairement, à l'aide du développement limité de f en a

$$\det J(f)(a) = \lim_{t\to 0} \frac{[f(a+ty_1)-f(a), \ldots, f(a+ty_n)-f(a)]}{[ty_1, \ldots, ty_n]}$$

que l'on peut interpréter comme rapport de volumes "infinitésimaux", au voisinage de a et de $f(a)$. Il ne sera donc pas étonnant de voir apparaître des valeurs absolues de déterminants Jacobiens lorsqu'on fera des changements de variables dans les intégrales multiples.

Pour $f : U \to \mathbb{E}_n$ différentiable en a, la non nullité du déterminant Jacobien de f en a signifie évidemment que df_a est un automorphisme de $\mathbb{E}_n$.

18-1.2.4 Gradient d'une fonction numérique sur un espace euclidien

Si $f : U \to \mathbb{R}$ est une fonction numérique différentiable en a, la différentielle df_a est une forme linéaire sur $\mathbb{E}_n$. Si $\mathcal{B}$ est une base de $\mathbb{E}_n$, la ligne représentant cette forme linéaire dans cette base est

$$(D_1 f(a), \ldots, D_n f(a)) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x}(a)\right)$$

ce qui est un autre moyen d'écrire

$$df_a = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, dx_j$$

Si $(\mathbb{E}_n, \langle\, , \rangle)$ est un espace euclidien, on sait qu'une forme linéaire $l \in \mathbb{E}_n^*$ s'écrit de manière unique $\langle z, \bullet\rangle$, avec $z \in \mathbb{E}_n$. Ceci amène à la définition

DÉFINITION 18-1.17 *Si U est un ouvert d'un espace euclidien $(\mathbb{E}_n, \langle\, , \rangle)$ et $f : U \to \mathbb{R}$ est différentiable en a, on appelle (vecteur) gradient de f en a le vecteur* $\operatorname{grad} f(a)$ *défini par*

$$\forall h \in \mathbb{E}_n \quad df_a(h) = \langle \operatorname{grad} f(a), h\rangle$$

Le développement limité de f en a s'écrit donc

$$f(a+h) = f(a) + \langle \operatorname{grad} f(a), h\rangle + \|h\|\,\varepsilon(h)$$

Les coordonnées du gradient s'expriment aisément dans une base orthonormale de $\mathbb{E}_n$: si $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ est une telle base, pour $h = \sum_{i=1}^{n} h_i e_i$, on a

$$df_a(h) = \sum_{j=1}^{n} h_j D_j f(a) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) = \left\langle \sum_{i=1}^{n} h_i e_i, \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} e_j \right\rangle$$

puisque la base est orthonormale. Donc

$$\operatorname{grad} f(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} e_j$$

(si la base n'était pas orthonormale, la matrice du produit scalaire dans cette base interviendrait dans le calcul). Il importe de bien comprendre que ce vecteur ne dépend pas de la base orthonormale choisie.

18-1.3 Condition suffisante de différentiabilité. Caractérisation des fonctions C^1

Pour le moment, nous savons écrire la matrice d'une différentielle, en utilisant les dérivées partielles ... sous réserve d'existence de cette différentielle ! Les résultats qui suivent permettront d'assurer cette existence.

18-1.3.1 Condition suffisante de différentiabilité

THÉORÈME 18-1.18 Soit $f : U \to \mathbb{F}_p$, avec U ouvert de $\mathbb{E}_n$. On suppose qu'il existe une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ de $\mathbb{E}_n$ telle que, au voisinage de a, f possède des dérivées partielles par rapport à chacune des coordonnées dans $\mathcal{B}$, les fonctions (définies au voisinage de a)

$$x \mapsto D_i f(x) = \frac{\partial f}{\partial x_i}(x)$$

étant toutes continues en a. Alors f est différentiable en a, avec bien sûr

$$df_a\left(\sum_{j=1}^{n} h_i e_i\right) = \sum_{j=1}^{n} h_j D_j f(a)$$

Démonstration : En remplaçant f par la fonction $x \mapsto f(a+x)$ (définie sur l'ouvert $-a+U$), on peut se ramener à $a = 0_{\mathbb{E}_n}$. En remplaçant ensuite la fonction f par la fonction

$$x = \sum_{j=1}^{n} x_j e_j \mapsto f(x) - f(0_{\mathbb{E}_n}) - \sum_{j=1}^{n} x_j D_j f(0_{\mathbb{E}_n})$$

on peut aussi supposer que toutes les dérivées partielles de f sont nulles à l'origine. Prouver la différentiabilité de f en $0_{\mathbb{E}_n}$ revient alors à démontrer que

$$f(h) \underset{h \to 0_{\mathbb{E}_n}}{=} o(h)$$

On travaille dans $\mathbb{E}_n$ avec la norme

$$\left\| \sum_{j=1}^{n} h_j e_j \right\| = \max_{1 \leqslant j \leqslant n} |h_j|$$

Si $\varepsilon > 0$ est fixé, par continuité des dérivées partielles en $0_{\mathbb{E}_n}$, on peut trouver un réel $\alpha > 0$ tel que

$$\forall y \in \mathbb{E}_n \quad \|y\| \leqslant \alpha \Rightarrow \forall j \in \{1, \dots, n\} \quad \|D_j f(x)\| \leqslant \varepsilon$$

On a alors, pour $h = \sum_{j=1}^{n} h_j e_j$ vérifiant $\|h\| \leqslant \alpha$

$$f(h) = \sum_{k=0}^{n-2} \left[f\left(\sum_{j=1}^{n-k} h_j e_j\right) - f\left(\sum_{j=1}^{n-k-1} h_j e_j\right) \right] + f(h_1 e_1)$$

(pour se déplacer de $0_{\mathbb{E}_n}$ à h, on passe par les points intermédiaires $h_1 e_1$, puis $h_1 e_1 + h_2 e_2$ etc... en suivant des parallèles aux axes de coordonnées, ce qui fait apparaître dans l'expression précédente des accroissements d'applications partielles construites à partir de f). La quantité $f(h)$ apparaît donc comme une somme de n termes. On montre que chacun de ces termes est majoré par $\varepsilon \|h\|$. En effet, pour $0 \leqslant k \leqslant n-2$

$$f\left(\sum_{j=1}^{n-k} h_j e_j\right) - f\left(\sum_{j=1}^{n-k-1} h_j e_j\right)$$

est l'accroissement sur le segment $[0, h_{n-k}]$ de la fonction

$$t \mapsto \varphi(t) = f\left(\sum_{j=1}^{n-k-1} h_j e_j + t e_{n-k}\right)$$

qui est dérivable sur ce segment, de dérivée

$$\varphi'(t) = D_{n-k} f\left(\sum_{j=1}^{n-k-1} h_j e_j + t e_{n-k}\right)$$

(par définition de la dérivée partielle). Si $\|h\| \leqslant \alpha$, cette dérivée est majorée en norme par ε sur le segment $[0, h_{n-k}]$, et par inégalité des accroissements finis

$$\left\| f\left(\sum_{j=1}^{n-k} h_j e_j\right) - f\left(\sum_{j=1}^{n-k-1} h_j e_j\right) \right\| \leqslant \varepsilon |h_{n-k}| \leqslant \varepsilon \|h\|$$

Le terme $f(h_1 e_1)$ se majorant de la même façon, on a finalement par inégalité triangulaire

$$\|h\| \leqslant \alpha \Rightarrow \|f(h)\| \leqslant n\varepsilon \|h\|$$

ce qui prouve le théorème. ■

EXERCICE 18-1.19 Soit $\mathcal{A}$ un espace affine euclidien de dimension n. Si $\varphi : \mathbf{R}^+ \to \mathbf{R}$ est une fonction de classe C^1, montrer que

$$\psi : \mathcal{A} \to \mathbf{R} \quad M \mapsto \varphi\left(\left\|\overrightarrow{OM}\right\|\right)$$

est différentiable en tout point de $\mathcal{A} - \{O\}$ et calculer son gradient en un point M.

Si on travaille dans un repère orthonormé d'origine O et de vecteurs de base $(\overrightarrow{e_i})_{1 \leqslant i \leqslant n}$, et si M a pour coordonnées $(x_1, \ldots, x_n)$, on a

$$\psi(M) = \varphi\left(\sqrt{x_1^2 + \cdots + x_n^2}\right)$$

En tout point de $\mathcal{A} - \{O\}$, cette fonction possède des dérivées partielles par rapport à chaque coordonnée, et on a

$$\frac{\partial \psi}{\partial x_i}(M) = \varphi'\left(\sqrt{x_1^2 + \cdots + x_n^2}\right) \frac{x_i}{\sqrt{x_1^2 + \cdots + x_n^2}}$$

(dérivée d'une fonction composée). Ces fonctions sont continues sur $\mathcal{A} - \{O\}$, et ψ est différentiable en tout point de cet ouvert, avec

$$\operatorname{grad} \psi(M) = \sum_{i=1}^{n} \frac{\partial \psi}{\partial x_i}(M)\, \overrightarrow{e_i} = \varphi'\left(\left\|\overrightarrow{OM}\right\|\right) \frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|}$$

Dans le cas de la dimension 2 et avec l'utilisation de coordonnées polaires d'origine O, en notant comme d'habitude

$$\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|} = \overrightarrow{u}(\theta)$$

et $r = \left\|\overrightarrow{OM}\right\|$, on aura

$$\operatorname{grad} \varphi(r) = \varphi'(r)\, \overrightarrow{u}(\theta)$$

Nous avons obtenu ce résultat par un calcul de dérivées partielles. On peut préférer le calcul direct de développement limité :

$$\left\|\overrightarrow{OM} + \vec{h}\right\| = \sqrt{\left\|\overrightarrow{OM} + \vec{h}\right\|^2} = \sqrt{\overrightarrow{OM}^2 + 2\vec{h}.\overrightarrow{OM} + \vec{h}^2}$$

$$= \left\|\overrightarrow{OM}\right\| \sqrt{1 + 2\vec{h}.\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|^2} + o\left(\left\|\vec{h}\right\|\right)}$$

ce qui donne, puisque par inégalité de Schwarz $\vec{h}.\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|^2} = O\left(\left\|\vec{h}\right\|\right)$,

$$\left\|\overrightarrow{OM} + \vec{h}\right\| = \left\|\overrightarrow{OM}\right\| + \vec{h}.\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|} + o\left(\left\|\vec{h}\right\|\right)$$

soit finalement

$$\varphi\left(M+\vec{h}\right)=\varphi\left(\left\|\overrightarrow{OM}\right\|+\vec{h}.\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|}+o\left(\left\|\vec{h}\right\|\right)\right)$$

$$=\varphi\left(\left\|\overrightarrow{OM}\right\|\right)+\varphi'\left(\left\|\overrightarrow{OM}\right\|\right)\frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|}.\vec{h}+o\left(\left\|\vec{h}\right\|\right)$$

ce qui redonne le gradient de φ en M.

18-1.3.2 Caractérisation des fonctions de classe $\mathcal{C}^1$

A la section **18-1.1.3**, nous avons donné une définition "intrinsèque" (indépendante du choix d'une base) pour les fonctions de classe $\mathcal{C}^1$. Nous donnons ici une caractérisation à l'aide des dérivées partielles dans une base.

THÉORÈME 18-1.20 Une fonction $f : U \to \mathbb{F}_p$ est de classe $\mathcal{C}^1$ sur l'ouvert U de $\mathbb{E}_n$ si et seulement s'il existe une base $\mathcal{B}$ de $\mathbb{E}_n$ telle que f admette en tout point de U des dérivées partielles par rapport à chaque coordonnée dans $\mathcal{B}$, les n applications

$$D_i f : U \to \mathbb{F}_p \text{ **définies par** } a \mapsto D_i f(a) = \frac{\partial f}{\partial x_i}(a)$$

(avec $1 \leqslant i \leqslant n$) étant continues sur U. Toutes les bases de $\mathbb{E}_n$ vérifient alors cette propriété.

Démonstration : Si $f \in \mathcal{C}^1(U, \mathbb{F}_p)$, et si $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ est une base de $\mathbb{E}_n$, la dérivée partielle de f en un point a selon la $i^{\text{ème}}$ coordonnée dans $\mathcal{B}$ n'est autre que

$$D_i f(a) = D_{e_i} f(a)$$

et dépend évidemment continûment de a.

Réciproquement, si dans une base $\mathcal{B}$ les dérivées partielles existent en tout point et sont continues, la fonction f est différentiable en tout point de U, et

$$\forall a \in U \quad \forall x \in \mathbb{E}_n \quad df_a(x) = D_x f(a) = \sum_{i=1}^{n} x_i D_i f(a) \text{ si } x = \sum_{i=1}^{n} x_i e_i$$

La fonction $a \mapsto D_x f(a)$ est donc définie et continue sur U, comme combinaison linéaire de fonctions continues. La fonction f est bien de classe $\mathcal{C}^1$ sur U. ■

COROLLAIRE 18-1.21 Une application $f : U \to \mathbb{F}_p$ est de classe $\mathcal{C}^1$ sur U si et seulement si f est différentiable en tout point de U et si l'application

$$U \to \mathcal{L}(\mathbb{E}_n, \mathbb{F}_p) \quad \text{**définie par**} \quad a \mapsto df_a$$

est continue sur U. On dit alors aussi que f est continûment différentiable[5] sur U.

Démonstration : Si f est de classe C^1, elle est différentiable en tout point de U. L'espace $\mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)$ étant de dimension finie, on le munit d'une norme quelconque. En fixant des bases $\mathcal{B}$ et $\mathcal{B}'$ de $\mathbb{E}_n$ et $\mathbb{F}_p$, dire que l'application $a \mapsto df_a$ est continue revient à montrer que $a \mapsto M(df_a, \mathcal{B}, \mathcal{B}')$ est continue de U dans $\mathcal{M}_{p,n}(\mathbb{R})$. Avec les notations de la section **18-1.2.3**, cette matrice s'écrit

$$J(f)(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_j}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \dfrac{\partial f_i}{\partial x_1}(a) & \cdots & \dfrac{\partial f_i}{\partial x_j}(a) & \cdots & \dfrac{\partial f_i}{\partial x_n}(a) \\ \vdots & & \vdots & & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \cdots & \dfrac{\partial f_p}{\partial x_j}(a) & \cdots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

et est bien fonction continue de a si f est C^1.

Si réciproquement $a \mapsto df_a$ est continue, les coefficients de la matrice jacobienne sont des fonctions continues, ce qui traduit exactement le caractère C^1 de f. On pourrait dire aussi que, pour $x \in \mathbb{E}_n$ quelconque

$$a \mapsto D_x f(a) = df_a(x)$$

est continue sur U. ■

COROLLAIRE 18-1.22 $C^1(U, \mathbb{F}_p) \subset C^0(U, \mathbb{F}_p)$.

C'est heureux pour la cohérence des notations !

18-1.3.3 Exemples

EXERCICE 18-1.23 Montrer que l'application

$$\varphi : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R} \quad \text{définie par } M \mapsto \det M$$

est de classe C^1 sur $\mathcal{M}_n(\mathbb{R})$, et montrer que, pour $H \in \mathcal{M}_n(\mathbb{R})$

$$d\varphi_M(H) = \operatorname{trace}\left({}^t\operatorname{com} M\right) H$$

On peut faire un calcul de dérivées partielles dans la base canonique. On peut aussi calculer la dérivée de φ en une matrice M inversible selon H, en utilisant la notion de polynôme caractéristique.

[5] Si l'espace de départ est de dimension infinie, la notion de dérivée partielle n'est plus utilisable. C'est alors cette propriété qu'on utiliserait pour définir les fonctions de classe C^1 (en munissant évidemment l'espace $\mathcal{L}_c(\mathbb{E}, \mathbb{F})$ de la norme des applications linéaires continues).

EXERCICE 18-1.24 On sait que $\mathcal{GL}_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$. Montrer que l'application

$$M \mapsto M^{-1}$$

est de classe C^1 sur $\mathcal{GL}_n(\mathbb{R})$ et calculer sa différentielle. L'utilisation des dérivées partielles ne semble pas ici une bonne idée ! On pourra calculer la différentielle en dérivant l'application définie au voisinage de 0 par $\psi(t) = (M+tH)^{-1}(M+tH)$. On peut aussi faire directement un développement limité en utilisant les résultats de la section **9-3.4**.

18-2 Calcul des différentielles

18-2.1 Linéarité

THÉORÈME 18-2.1 Si f et $g : U \to \mathbb{F}_p$ sont différentiables en a, il en est de même de toute combinaison linéaire de f et g, et on a

$$d(f+\lambda g)_a = df_a + \lambda dg_a$$

Avec un choix de bases de $\mathbb{E}_n$ et $\mathbb{F}_p$, on aura

$$J(f+\lambda g)(a) = J(f)(a) + \lambda J(g)(a)$$

Ce résultat est évidemment à rapprocher de celui de la section **18-1.1.3**. De même, le théorème qui suit n'est qu'un cas particulier du théorème de différentiation d'une fonction composée que nous évoquons à la section **18-2.3** :

THÉORÈME 18-2.2 Si $f : U \to \mathbb{F}_p$ est différentiable en a et si $v : \mathbb{F}_p \to \mathbb{G}_q$ est une application linéaire (continue !), la composée $v \circ f$ est différentiable en a avec

$$d(v \circ f)_a = v \circ df_a$$

Démonstration : Il suffit de composer l'égalité

$$f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon(h)$$

avec l'application linéaire v pour obtenir

$$v \circ f(a+h) = v \circ f(a) + v \circ df_a(h) + \|h\| v(\varepsilon(h))$$

ce qui prouve le résultat, puisque $\lim_{h \to 0_{\mathbb{E}_n}} v(\varepsilon(h)) = 0_{\mathbb{G}_q}$. ■

18-2.2 Produit. Inverse d'une fonction numérique

THÉORÈME 18-2.3 Soient $f : U \to \mathbb{F}_p$ et $g : U \to \mathbb{G}_q$ deux fonctions différentiables en a. Si $B : \mathbb{F}_p \times \mathbb{G}_q \to \mathbb{H}_r$ est une application bilinéaire (continue !), l'application

$$\Phi = B(f,g) : U \to \mathbb{H}_q \text{ définie par } \Phi(x) = B(f(x), g(x))$$

est différentiable en a, et

$$\forall h \in \mathbb{E}_n \quad d\Phi_a(h) = B(df_a(h), g(a)) + B(f(a), dg_a(h))$$

Démonstration : Ici encore, ce théorème se trouve être un cas particulier du théorème de différentiation d'une fonction composée. La démonstration directe est simple. Elle consiste à utiliser les développements limités de f et g. Pour $\|h\| \leqslant \alpha$, on a en effet

$$B(f(a+h), g(a+h)) = \\ B(f(a) + df_a(h) + \|h\| \varepsilon_1(h), g(a) + dg_a(h) + \|h\| \varepsilon_2(h))$$

qu'on développe par bilinéarité. On obtient

$$B(f(a+h), g(a+h)) = \\ B(f(a), g(a)) + B(df_a(h), g(a)) + B(f(a), dg_a(h)) + \psi(h)$$

où $\psi(h)$ est somme de 6 termes qui sont tous des $o(h)$, à cause de l'existence d'une constante $K > 0$ telle que

$$\forall x \in \mathbb{F}_p \quad \forall y \in \mathbb{G}_q \quad \|B(x,y)\| \leqslant K \|x\| \|y\|$$

inégalité qui traduit la continuité de B. Comme

$$h \mapsto B(df_a(h), g(a)) + B(f(a), dg_a(h))$$

est clairement linéaire (continue), le résultat est démontré. ■

Par exemple, si u et v sont deux fonctions numériques différentiables en a définies sur un espace euclidien, on a

$$\operatorname{grad}(uv)(a) = u(a) \operatorname{grad} v(a) + v(a) \operatorname{grad} u(a)$$

EXERCICE 18-2.4 Si u est une fonction numérique différentiable en a, avec $u(a) \neq 0$, montrer que

$$x \mapsto \frac{1}{u(x)}$$

est définie au voisinage de a, différentiable en a avec

$$\operatorname{grad} \frac{1}{u}(a) = -\frac{1}{u^2(a)} \operatorname{grad} u(a)$$

Si u est de classe C^1 au voisinage de a, il en est alors de même de $\frac{1}{u}$.

18-2.3 Différentielle d'une composée

Le théorème fondamental est celui de la section **18-2.3.3**. Il contient comme nous l'avons déjà dit les résultats des sections précédentes. Nous allons cependant commencer par l'étude d'un cas particulier, celui de la dérivation d'une fonction composée, qui est essentiel.

18-2.3.1 Dérivée d'une fonction composée

THÉORÈME 18-2.5 Soit I un intervalle de $\mathbb{R}$, et $\varphi : I \to \mathbb{E}_n$ dérivable en $t_0 \in I$. Soit U un ouvert de $\mathbb{E}_n$ contenant $\varphi(I)$ et $f : U \to \mathbb{F}_p$ différentiable en $a = \varphi(t_0)$. La fonction $f \circ \varphi : I \to \mathbb{F}_p \quad t \mapsto f(\varphi(t))$ est alors dérivable en t_0, de dérivée

$$(f \circ \varphi)'(t_0) = df_{\varphi(t_0)}(\varphi'(t_0))$$

Démonstration : On suppose t_0 intérieur à I (le cas d'une extrémité se traiterait de façon analogue). On a alors

$$\exists\, \alpha > 0 \quad |h| \leqslant \alpha \Rightarrow \varphi(t_0 + h) = \varphi(t_0) + h\varphi'(t_0) + h\varepsilon(h)$$

avec $\lim\limits_{h \to 0} \varepsilon(h) = 0_{\mathbb{E}_n}$. Comme f est différentiable en $a = \varphi(t_0)$, on a, pour $k \in \mathbb{E}_n$ avec $a + k \in U$,

$$f(a+k) = f(a) + df_a(k) + \|k\|\, \varepsilon_1(k)$$

avec $\lim\limits_{0_{\mathbb{E}_n}} \varepsilon_1 = 0_{\mathbb{F}_p}$. On en déduit, pour $|h| \leqslant \alpha$, compte tenu de la linéarité de df_a

$$f(\varphi(t_0 + h)) = f(\varphi(t_0)) + h df_a(\varphi'(t_0)) + |h|\, \psi(h)$$

avec pour $h \neq 0$,

$$\psi(h) = \frac{h}{|h|}.df_a(\varepsilon(h)) + \|\varphi'(t_0) + \varepsilon(h)\| .\varepsilon_1(h\varphi'(t_0) + h\varepsilon(h))$$

qui tend clairement vers $0_{\mathbb{E}_n}$ pour $h \to 0$. ■

Remarquons que la proposition **18-1**.10 est un cas particulier de ce résultat. Si on veut utiliser les dérivées partielles, on aura par exemple le

COROLLAIRE 18-2.6 Si $\varphi : I \to \mathbb{R}^n \quad t \mapsto (u_1(t), \ldots, u_n(t))$ est une fonction de classe C^1 définie sur un intervalle I de $\mathbb{R}$ et si f est une fonction de classe C^1 sur l'ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, avec $\varphi(I) \subset U$, alors la fonction

$$F : t \mapsto f(u_1(t), \ldots, u_n(t))$$

est de classe C^1 sur I, avec

$$\forall\, t \in I \quad F'(t) = \frac{d}{dt}\left(f(u_1(t), \ldots, u_n(t))\right) = \sum_{i=1}^{n} u_i'(t) \frac{\partial f}{\partial x_i}(u_1(t), \ldots, u_n(t))$$

Démonstration : La dérivabilité de F et la formule donnant $F'(t)$ s'obtiennent par application directe du théorème précédent. L'application $t \mapsto F'(t)$ est continue, comme somme de composées de fonctions continues. ■

On pourra en particulier appliquer ce corollaire pour faire des calculs de dérivées partielles (puisqu'il s'agit alors, en étudiant des applications partielles, de dériver des fonctions d'**une** variable réelle). Par exemple, si $f : \mathbb{R}^2 \to \mathbb{R}$ $(x,y) \mapsto f(x,y)$ est une fonction de classe $\mathcal{C}^1$, on aura

$$\frac{\partial}{\partial u}\left[f\left(u^2 v, e^u \sin v\right)\right] = 2uv \frac{\partial f}{\partial x}\left(u^2 v, e^u \sin v\right) + e^u \sin v \frac{\partial f}{\partial y}\left(u^2 v, e^u \sin v\right)$$

18-2.3.2 Interprétations géométriques

Tangente à l'image d'un arc paramétré :

Supposons que $\gamma : I \ni t \mapsto M(t) \in U$ soit un arc paramétré de classe $\mathcal{C}^1$ dont le support est inclus dans U. Si γ est **régulier** , c'est-à-dire

$$\forall t \in I \quad \vec{M}'(t) \neq \vec{0}$$

on sait que le vecteur dérivé $\vec{M}'(t)$ dirige la tangente à γ au point de paramètre t. Pour $f : \mathbb{R}^n \supset U \to \mathbb{R}^p$ de classe $\mathcal{C}^1$ (qu'on peut considérer comme une transformation géométrique $M \mapsto f(M)$, avec le plus souvent pour nous $n = p = 2$ ou 3) l'arc image par cette transformation

$$f(\gamma) : I \ni t \mapsto N(t) = f(M(t)) \in \mathbb{R}^p$$

est alors de classe $\mathcal{C}^1$, avec

$$\forall t \in I \quad \vec{N}'(t) = df_{M(t)}\left(\vec{M}'(t)\right)$$

Si ce vecteur est non nul en un point t_0 (ce qui est certainement le cas si γ est régulier et $df_{M(t_0)}$ est injective, en particulier si $n = p$ et $df_{M(t_0)} \in \mathcal{GL}(\mathbb{R}^n)$), alors l'arc image est régulier au point de paramètre t_0, et le vecteur tangent à $f(\gamma)$ en t_0 est image par $df_{M(t_0)}$ du vecteur tangent à γ en t_0. Il en résulte que la tangente à $f(\gamma)$ en $N(t_0) = N_0$

$$T_0' = N_0 + \mathbb{R}\vec{N}'(t_0)$$

est image de la tangente à γ en $M(t_0) = M_0$

$$T_0 = M_0 + \mathbb{R}\vec{M}'(t_0)$$

par l'application affine tangente (figure **18.1**)

$$\mathbb{R}^n \to \mathbb{R}^p \quad M \mapsto N = N_0 + df_{M_0}\left(\overrightarrow{M_0 M}\right)$$

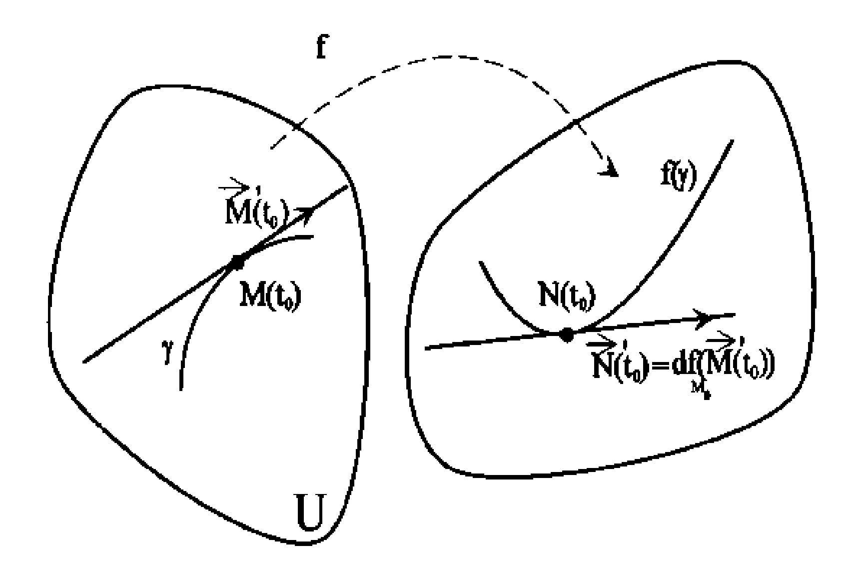


Figure **18.1** – *Tangente à l'image d'un arc paramétré*

EXERCICE 18-2.7 Si $\mathcal{E}_n$ est un espace affine euclidien et O est un point de $\mathcal{E}_n$, on appelle inversion de pôle O et de puissance $k \in \mathbb{R}^*$ l'application de $\mathcal{E}_n - \{O\}$ dans lui-même définie par

$$i : M \mapsto P \quad \text{avec} \quad O, M \text{ et } P \text{ alignés et } \overline{OM}.\overline{OP} = k$$

Montrer que

$$\overrightarrow{OP} = k \frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|^2}$$

En déduire que i est de classe C^1 sur $\mathcal{E}_n - \{O\}$. A l'aide d'un calcul intrinsèque inspiré de celui de l'exemple **18-1.19**, interpréter géométriquement la différentielle de i en un point $M_0 \in \mathcal{E}_n - \{O\}$. En déduire que l'inversion est une transformation conforme (conservant les angles), c'est-à-dire que, si

$$\gamma_1 : t \mapsto M(t) \text{ et } \gamma_2 : t \mapsto N(t)$$

sont deux arcs paramétrés réguliers sécants en M_0 pour $t = 0$, (les tangentes en ce point étant les droites notées T_0 et S_0), leurs images par l'inversion

$$i(\gamma_1) : t \mapsto P(t) \text{ et } i(\gamma_2) : t \mapsto Q(t)$$

ont des tangentes T_0' et S_0' en $P_0 = i(M_0)$ qui vérifient

$$(\widehat{T_0, S_0}) = (\widehat{T_0', S_0'})$$

les angles considérés ici étant des angles (non orientés!) de droites : si $T_0 = M_0 + \mathbb{R}\vec{u}$ et $S_0 = M_0 + \mathbb{R}\vec{v}$, $(\widehat{T_0, S_0})$ est l'unique réel θ de $[0, \pi]$ vérifiant

$$\cos\theta = \frac{\vec{u}.\vec{v}}{\|\vec{u}\| \|\vec{v}\|}$$

Plan tangent à une surface d'équation $z = f(x, y)$

Si U est un ouvert de $\mathbb{R}^2$, on peut, comme dans le cas d'une fonction d'une variable réelle, envisager une représentation géométrique de $f : U \to \mathbb{R}$: dans un espace affine $\mathcal{A}_3$ de dimension 3 rapporté à un repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$, on considèrera

$$(S) : \left\{M \in \mathcal{A}_3 \mid \exists\, (x, y) \in U \quad \overrightarrow{OM} = x\vec{\imath} + y\vec{\jmath} + f(x, y)\,\vec{k}\right\}$$

(S) est donc l'ensemble des points de coordonnées $M(x, y, f(x, y))$ avec $(x, y) \in U$. C'est une "surface" d'équation $z = f(x, y)$ dans le repère considéré.

Supposons f de classe C^1 dans U et considérons un point

$$M_0(x_0, y_0, z_0 = f(x_0, y_0)) \in (S)$$

Soit $\gamma : t \mapsto M(t)$ un arc paramétré de classe C^1 défini sur un intervalle I voisinage de 0 dans $\mathbb{R}$. On suppose que cet arc est tracé sur (S) et vérifie $M(0) = M_0$: il existe donc des fonctions numériques C^1 notées $t \mapsto x(t)$ et $t \mapsto y(t)$ vérifiant

$$\forall t \in I \quad (x(t), y(t)) \in U$$

et

$$\forall t \in I \quad \overrightarrow{OM(t)} = x(t)\,\vec{\imath} + y(t)\,\vec{\jmath} + f(x(t), y(t))\,\vec{k}$$

(avec $x(0) = x_0$ et $y(0) = y_0$). Par application du corollaire **18-2**.6, on a

$$\overrightarrow{M'}(0) = x'(0)\,\vec{\imath} + y'(0)\,\vec{\jmath} + \left[x'(0)\,\frac{\partial f}{\partial x}(x_0, y_0) + y'(0)\,\frac{\partial f}{\partial y}(x_0, y_0)\right]\vec{k}$$

L'arc γ est donc régulier en $t = 0$ si et seulement si $(x'(0), y'(0)) \neq (0, 0)$. On sait alors que la tangente à γ en M_0 de paramètre 0 est la droite

$$\mathcal{D}_0 = M_0 + \mathbb{R}.\overrightarrow{M'}(0)$$

Le calcul précédent montre que le vecteur $\overrightarrow{M'}(0)$ est nécessairement dans le plan vectoriel d'équation, dans la base $\left(\vec{\imath}, \vec{\jmath}, \vec{k}\right)$:

$$Z = \frac{\partial f}{\partial x}(x_0, y_0)\,X + \frac{\partial f}{\partial y}(x_0, y_0)\,Y$$

Réciproquement, tout vecteur non nul de ce plan

$$\vec{u} = \alpha\vec{\imath} + \beta\vec{\jmath} + \left[\frac{\partial f}{\partial x}(x_0, y_0)\,\alpha + \frac{\partial f}{\partial y}(x_0, y_0)\,\beta\right]\vec{k}$$

est tangent en M_0 à un arc γ particulier, par exemple celui obtenu pour

$$x(t) = x_0 + \alpha \text{ et } y(t) = y_0 + \beta t$$

On en déduit la

PROPOSITION 18-2.8 Si $f : U \to \mathbb{R}$ est une application de classe C^1, la réunion des tangentes aux arcs réguliers tracés sur la surface d'équation $z = f(x, y)$ en un point $M_0(x_0, y_0, z_0 = f(x_0, y_0))$ forme un plan affine, appelé plan tangent à la surface en M_0. Ce plan T_0 a pour équation

$$Z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(X - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(Y - y_0)$$

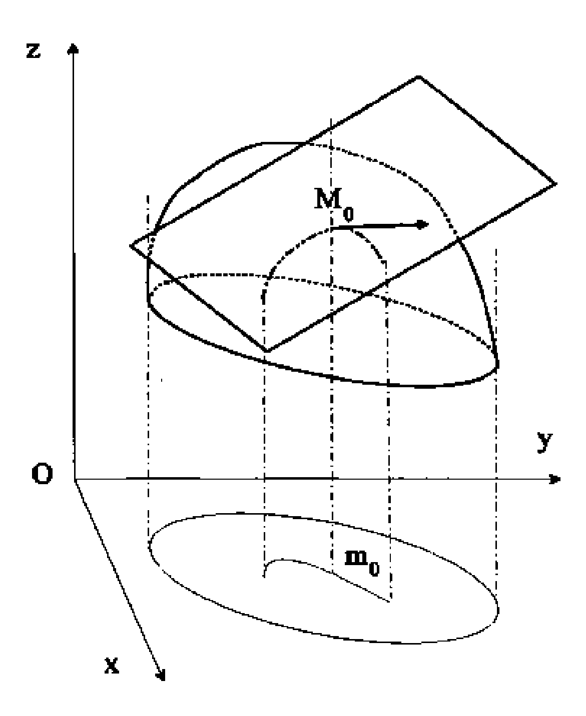


Figure **18.2** – *Plan tangent à une surface d'équation* $z = f(x, y)$

18-2.3.3 Différentielle d'une composée

Nous arrivons à l'énoncé du théorème général qui englobe tous les résultats particuliers obtenus précédemment :

THÉORÈME 18-2.9 Soient U **et** V **des ouverts respectifs de** $\mathbb{E}_n$ **et** $\mathbb{F}_p$, $f : U \to V$ **différentiable en** $a \in U$, $g : V \to \mathbb{G}_q$ **différentiable en** $b = f(a)$. **La composée** $g \circ f$ **est alors différentiable en** a, **et l'on a**

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

Démonstration : Il suffit simplement de composer les développements limités. Remarquons d'abord que $dg_{f(a)} \circ df_a$ est une application linéaire de $\mathbb{E}_n$ vers $\mathbb{G}_q$. Comme f est différentiable en a, on peut trouver un $\alpha > 0$ tel que, pour $h \in \mathbb{E}_n$

$$\|h\| \leqslant \alpha \Rightarrow a + h \in U \text{ et } f(a+h) = f(a) + df_a(h) + \|h\| \varepsilon(h)$$

avec $\lim\limits_{h \to 0_{\mathbb{E}_n}} \varepsilon(h) = 0_{\mathbb{F}_p}$. On a de même, pour $b + k \in V$

$$g(b+k) = g(b) + dg_b(k) + \|k\| \varepsilon_1(k)$$

avec $\lim\limits_{k \to 0_{\mathbb{F}_p}} \varepsilon_1(k) = 0_{\mathbb{G}_q}$. On a alors, pour $h \in \mathbb{E}_n$ et $\|h\| \leqslant \alpha$

$$\begin{aligned} g(f(a+h)) = g(f(a)) &+ dg_b(df_a(h) + \|h\| \varepsilon(h)) \\ &+ \|df_a(h) + \|h\| \varepsilon(h)\| \varepsilon_1(df_a(h) + \|h\| \varepsilon(h)) \end{aligned}$$

ce qui s'écrit, compte tenu de la linéarité de dg_b

$$g(f(a+h)) = g(f(a)) + dg_b(df_a(h)) + \|h\|\,\phi(h)$$

avec, pour $h \neq 0_{\mathbb{E}_n}$

$$\phi(h) = dg_b(\varepsilon(h)) + \left\|\frac{df_a(h)}{\|h\|} + \varepsilon(h)\right\| \varepsilon_1(df_a(h) + \|h\|\,\varepsilon(h))$$

fonction qui tend vers $0_{\mathbb{G}_q}$ pour $h \to 0_{\mathbb{E}_n}$, puisque le rapport $\dfrac{df_a(h)}{\|h\|}$ reste borné, par continuité de df_a. ■

Si l'on ramène les espaces $\mathbb{E}_n$, $\mathbb{F}_p$ et $\mathbb{G}_q$ à des bases $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$, les différentielles sont caractérisées par les matrices jacobiennes $J(f)(a) \in \mathcal{M}_{p,n}(\mathbb{R})$ et $J(g)(f(a)) \in \mathcal{M}_{q,p}(\mathbb{R})$. On a alors immédiatement

COROLLAIRE 18-2.10 Sous les hypothèses du théorème précédent

$$J(g\circ f)(a) = J(g)(f(a)) \times J(f)(a)$$

COROLLAIRE 18-2.11 Si $f : U \to V$ et $g : V \to \mathbb{G}_q$ sont deux fonctions de classe $\mathcal{C}^1$, leur composée est de classe $\mathcal{C}^1$.

Démonstration : On fixe des bases des $\mathbb{E}_n$, $\mathbb{F}_p$ et $\mathbb{G}_q$. L'application

$$U \to \mathcal{M}_{p,n}(\mathbb{R}) \text{ définie par } a \mapsto J(f)(a)$$

est continue, puisque f est $\mathcal{C}^1$. De même l'application

$$U \to \mathcal{M}_{q,p}(\mathbb{R}) \text{ définie par } a \mapsto J(g)(f(a))$$

est continue, comme composée de $f : U \to V$ et de $V \to \mathcal{M}_{q,p}(\mathbb{R})$ définie par $y \mapsto J(g)(y)$. Par continuité du produit matriciel

$$\mathcal{M}_{q,p}(\mathbb{R}) \times \mathcal{M}_{p,n}(\mathbb{R}) \to \mathcal{M}_{q,n}(\mathbb{R})$$

on obtient la continuité des dérivées partielles de la fonction $g\circ f$. ■

Ces corollaires contiennent évidemment les résultats sur les calculs de dérivées partielles vus à la fin de la section **18-2.3.1** : par exemple, si

$$f : U \to \mathbb{R} \quad (x,y) \mapsto f(x,y)$$

est une fonction de classe $\mathcal{C}^1$ sur un ouvert de $\mathbb{R}^2$, et si $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ est l'application définie par $(r,\theta) \mapsto (r\cos\theta, r\sin\theta)$ (évidemment de classe $\mathcal{C}^1$ comme le montre le calcul de ses dérivées partielles) correspondant au passage en coordonnées polaires, l'application obtenue par changement de variable

$$F : \Phi^{-1}(U) \to \mathbb{R} \quad (r,\theta) \mapsto f(r\cos\theta, r\sin\theta)$$

est de classe C^1 sur l'ouvert $\Phi^{-1}(U)$ et les égalités

$$\left\{\begin{array}{l} \dfrac{\partial F}{\partial r}(r,\theta) = \cos\theta\dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \\ \dfrac{\partial F}{\partial \theta}(r,\theta) = -r\sin\theta\dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta\dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{array}\right.$$

peuvent effectivement s'écrire, en utilisant les matrices jacobiennes

$$\left(\frac{\partial F}{\partial r}, \frac{\partial F}{\partial \theta}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Application : expression du gradient en coordonnées polaires

Dans un plan affine eucidien rapporté à un repère orthonormé direct $(O, \vec{\imath}, \vec{\jmath})$, on considère un champ scalaire de classe C^1 défini sur un ouvert U :

$$M(x,y) \mapsto f(M) = f(x,y) \in \mathbb{R}$$

En tout point de U, on a

$$\overrightarrow{\text{grad}\, f}(M) = \frac{\partial f}{\partial x}(x,y)\,\vec{\imath} + \frac{\partial f}{\partial x}(x,y)\,\vec{\jmath}$$

Si (r,θ) est un système de coordonnées polaires du point $M \in U$, c'est-à-dire si

$$\overrightarrow{OM} = r\,\vec{u}(\theta)\,, \text{ avec } \vec{u}(\theta) = \cos\theta\,\vec{\imath} + \sin\theta\,\vec{\jmath}$$

on travaillera souvent avec le repère local $(M, \vec{u}(\theta), \vec{v}(\theta))$ avec

$$\vec{v}(\theta) = \vec{u}\left(\theta + \frac{\pi}{2}\right) = -\sin\theta\,\vec{\imath} + \cos\theta\,\vec{\jmath}$$

On a alors

$$\overrightarrow{\text{grad}\, f}(M) = \frac{\partial f}{\partial x}(x,y)(\cos\theta\,\vec{u} - \sin\theta\,\vec{v}) + \frac{\partial f}{\partial y}(x,y)(\sin\theta\,\vec{u} + \cos\theta\,\vec{v})$$

soit

$$\overrightarrow{\text{grad}\, f}(M) = \left[\frac{\partial f}{\partial x}(x,y)\cos\theta + \frac{\partial f}{\partial y}(x,y)\sin\theta\right]\vec{u} + \left[-\sin\theta\frac{\partial f}{\partial x}(x,y) + \cos\theta\frac{\partial f}{\partial y}\right]\vec{v}$$

Compte tenu des calculs effectués plus tôt, on aura, si M n'est pas l'origine du repère (c'est-à-dire si $r \neq 0$)

$$\boxed{\overrightarrow{\text{grad}\, f}(M) = \frac{\partial F}{\partial r}(r,\theta)\,\vec{u} + \frac{1}{r}\frac{\partial F}{\partial \theta}(r,\theta)\,\vec{v}}$$

Le terme en $\dfrac{1}{r}$ non défini à l'origine montre que, comme nous le verrons ultérieurement, le passage en coordonnées polaires n'est pas un "bon" changement de variables au voisinage de O.

18-2.3.4 Intérêt de la notation différentielle

Dans le calcul précédent, nous avons pris le soin de noter différemment les fonctions f et F. C'est conforme à la définition d'une fonction en Mathématiques. L'usage, en Physique notamment, est cependant de noter

$$f(M) = f(x,y) = f(r,\theta)$$

lorsque M est le point de coordonnées cartésiennes (x,y) et de coordonnées polaires (r,θ). La notation différentielle est alors un moyen commode de traduire "mécaniquement" le théorème de composition de la section précédente :

Lorsqu'on écrit

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

si f est effectivement fonction des variables x et y, cela doit être compris au sens défini juste après le théorème **18-1**.14. Par contre, si x et y sont considérées comme fonction de r et θ

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \Rightarrow \begin{cases} dx = \cos\theta.dr - r\sin\theta.d\theta \\ dy = \sin\theta.dr + r\cos\theta.d\theta \end{cases}$$

en remplaçant formellement dx et dy dans l'expression de df, on obtient

$$df = \left(\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right)dr + \left(\frac{\partial f}{\partial y}r\cos\theta - \frac{\partial f}{\partial x}r\sin\theta\right)d\theta$$

qui est bien l'expression de la différentielle de f considérée à présent comme fonction de r et θ.

Ce calcul formel est justifié par la formule donnant la matrice jacobienne d'une composée : lorsqu'on fait un changement de variables, remplacer dans l'égalité

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}dx_j$$

formellement dx_j par son expression à l'aide des nouvelles variables "indépendantes", revient exactement à faire le calcul du produit de deux matrices jacobiennes. C'est le cas notamment lorsque les x_i deviennent fonctions (C^1) d'une variable t, $t \mapsto x_i(t)$, où

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}dx_j = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_1(t),\ldots,x_n(t))\,x_i'(t)\,dt$$

redonne la formule de dérivation d'une fonction composée.

18-2.3.5 Exercice d'application

EXERCICE 18-2.12 Soit U un ouvert de $\mathbb{R}^2$ contenant un pavé compact

$$K = [a,b] \times [c,d]$$

et

$$f = U \to \mathbb{R} \quad (x,y) \mapsto f(x,y)$$

une application continue possédant une dérivée partielle $\dfrac{\partial f}{\partial y}$ continue sur U. Soit I un intervalle de $\mathbb{R}$, $u : I \to [a,b]$ et $v : I \to [c,d]$ deux applications de classe $\mathcal{C}^1$. On définit $F : I \to \mathbb{R}$ par

$$F(t) = \int_a^{u(t)} f(s, v(t))\, ds$$

Montrer que F est de classe $\mathcal{C}^1$ et calculer sa dérivée.

Indication : On définit Φ sur un pavé ouvert contenant K et inclus dans U par

$$\Phi(U,V) = \int_a^U f(s,V)\, ds$$

Cette application possède en tout point de son ouvert de définition des dérivées partielles

$$\frac{\partial \Phi}{\partial U}(U,V) = f(U,V) \text{ et } \frac{\partial \Phi}{\partial V}(U,V) = \int_a^U \frac{\partial f}{\partial y}(s,V)\, ds$$

ces fonctions étant continues (cela doit être prouvé avec précision pour $\dfrac{\partial \Phi}{\partial V}$). On en déduira facilement

$$\forall t \in I \quad F'(t) = f(u(t), v(t))\, u'(t) + v'(t) \int_a^{u(t)} \frac{\partial f}{\partial y}(s, v(t))\, ds$$

On obtient ainsi une généralisation du théorème de dérivation sous le signe intégrale. Par exemple, ce théorème pourrait s'appliquer pour dériver une fonction de la forme

$$F(t) = \int_0^t \sin(t-s)\, \varphi(s)\, ds$$

avec $\varphi : \mathbb{R} \to \mathbb{R}$ continue. Mais sur un exemple aussi simple, l'avantage resterait à qui connaîtrait ses formules de trigonométrie et penserait à écrire

$$F(t) = \sin t \int_0^t \cos s.\varphi(s)\, ds - \cos t \int_0^t \sin s.\varphi(s)\, ds$$

18-2.4 Différentielle d'une réciproque

Soient U et V des ouverts (non vides) respectifs des espaces normés $\mathbb{E}_n$ et $\mathbb{F}_p$, et f un homéomorphisme de U vers V. On démontre (et la démonstration n'est pas simple pour un résultat qu'on peut considérer comme "intuitif") que pour qu'un tel homéomorphisme existe, il est nécessaire que l'on ait $n = p$. Si on suppose f différentiable en un point a et que l'on souhaite la différentiabilité de f^{-1} en $f(a)$, cette condition sur les dimensions est alors simple à obtenir, puisque le théorème qui suit montre qu'alors df_a doit être un isomorphisme entre les espaces $\mathbb{E}_n$ et $\mathbb{F}_p$. Ce résultat généralise le théorème **8-4**.12 correspondant au cas $n = p = 1$.

THÉORÈME 18-2.13 Soit $f : U \to V$ un homéomorphisme d'un ouvert de $\mathbb{E}_n$ vers un ouvert de $\mathbb{F}_n$. On suppose f différentiable en un point $a \in U$. Pour que f^{-1} soit différentiable en $b = f(a)$, il faut et il suffit que df_a soit un isomorphisme entre $\mathbb{E}_n$ et $\mathbb{F}_n$, et on a alors

$$df_b^{-1} = (df_a)^{-1}$$

Démonstration : Supposons f^{-1} différentiable en b. Comme

$$f^{-1} \circ f = id_U \text{ et } f \circ f^{-1} = id_V$$

on obtient, par application du théorème de différentiation d'une composée

$$df_b^{-1} \circ df_a = id_{\mathbb{E}_n} \text{ et } df_a \circ df_b^{-1} = id_{\mathbb{F}_p}$$

ce qui montre bien que df_a est un isomorphisme de $\mathbb{E}_n$ vers $\mathbf{F}_p$ (donc $n = p$) et $df_b^{-1} = (df_a)^{-1}$.

Supposons réciproquement $l = df_a$ bijective, et notons $l' = l^{-1}$ ainsi que $g = f^{-1}$. Comme V est un ouvert, il existe un réel $\alpha > 0$ tel que

$$k \in \mathbb{F}_p \text{ et } \|k\| \leqslant \alpha \Rightarrow b + k \in V$$

Pour un tel k, il existe un unique vecteur $h(k) \in \mathbb{E}_n$ tel que

$$a + h(k) \in U \text{ et } f(a + h(k)) = b + k$$

(puisque f est supposée bijective de U vers V). On a alors

$$h(k) = g(b + k) - g(b)$$

et la continuité de g en b entraîne que

$$\lim_{k \to 0_{\mathbb{F}_p}} h(k) = 0_{\mathbb{E}_n}$$

Comme f est différentiable en a, on peut écrire

$$k = b + k - b = f(a + h(k)) - f(a) = l(h(k)) + \|h(k)\| \varepsilon(h(k))$$

où l'on sait que

$$\lim_{k \to 0_{\mathbb{F}_p}} \varepsilon(h(k)) = 0_{\mathbb{F}_p}$$

Comme l est bijective, on en déduit

$$g(b+k) - g(b) = h(k) = l'(k) - \|h(k)\| \, l'[\varepsilon(h(k))]$$

égalité qui nous donnera bien un développement limité de g au voisinage de b si l'on montre que $\|h(k)\| \underset{k \to 0_{\mathbb{F}_p}}{=} O(\|k\|)$. Ceci est conséquence de l'égalité précédente, puisqu'il existe un réel $\beta > 0$ tel que

$$\|k\| \leqslant \beta \Rightarrow \|l'[\varepsilon(h(k))]\| \leqslant \frac{1}{2}$$

l'inégalité triangualire donnant alors

$$\|k\| \leqslant \beta \Rightarrow \|h(k)\| \leqslant \|l'\| \, \|k\| + \frac{1}{2} \|h(k)\|$$

qui implique enfin $\|h(k)\| \leqslant 2 \|l'\| \, \|k\|$. ■

COROLLAIRE 18-2.14 Avec les hypothèses du théorème précédent, une fois fixées des bases de $\mathbb{E}_n$ et $\mathbb{F}_n$, on a

$$J(f^{-1})(b) = [J(f)(a)]^{-1}$$

COROLLAIRE 18-2.15 Si $f : U \to V$ est un homéomorphisme[6] de classe C^1 dont la différentielle est inversible en tout point de U, alors f^{-1} est de classe C^1 de V vers U.

Démonstration : Le théorème précédent entraîne la différentiabilité de f^{-1} en tout point de V. De plus, l'application

$$V \to \mathcal{M}_n(\mathbb{R}) : b \mapsto J(f^{-1})(b)$$

est continue comme composée de $f^{-1} : V \to U$, de l'application de U dans $\mathcal{GL}_n(\mathbb{R})$ définie par $a \mapsto J(f)(a)$ et de l'application $M \mapsto M^{-1}$ de $\mathcal{GL}_n(\mathbb{R})$ dans lui-même. ■

EXEMPLE 18-2.16 Reprenons l'exemple du passage en coordonnées polaires. Nous avons vu à la section **9-6.2.3** que l'application

$$\begin{aligned} \Phi :]0, +\infty[\times] - \pi, \pi[&\to \mathbb{R}^2 - \mathbb{R}^- \times \{0\} \\ (r, \theta) &\mapsto (x, y) = (r\cos\theta, r\sin\theta) \end{aligned}$$

est un homéomorphisme entre deux ouverts de $\mathbb{R}^2$. On a d'ailleurs dans ce cas une expression de Φ^{-1} :

$$\Phi^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, 2\arctan \frac{y}{x + \sqrt{x^2 + y^2}} \right)$$

[6]Nous verrons ultérieurement (théorème d'inversion globale) que cette hypothèse peut être considérablement affaiblie.

L'application Φ est évidemment de classe $\mathcal{C}^1$, sa matrice jacobienne (on travaille évidemment dans la base canonique de $\mathbb{R}^2$) étant

$$J(\Phi)(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Le déterminant de cette matrice vaut

$$\det J(\Phi)(r,\theta) = r$$

et ne s'annule en aucun point de l'ouvert de définition de Φ. Il en résulte que Φ^{-1} est différentiable en tout point de $\mathbb{R}^2 - \mathbb{R}^- \times \{0\}$ (elle est en fait de classe $\mathcal{C}^1$), et que la jacobienne de cette application en $(x,y) = (r\cos\theta, r\sin\theta)$ est tout simplement l'inverse de la matrice précédente :

$$J(\Phi^{-1})(x,y) = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

Si, comme le veut l'usage, nous notons $(x,y) \mapsto (r(x,y), \theta(x,y))$ l'application Φ^{-1}, remarquons que cette égalité matricielle permet d'évaluer les dérivées partielles des fonctions $(x,y) \mapsto r(x,y)$ et $(x,y) \mapsto \theta(x,y)$ sans qu'il soit nécessaire de connaître les formules explicites définissant ces fonctions : on a, avec la première ligne de cette matrice,

$$\frac{\partial r}{\partial x} = \cos\theta \quad \text{et} \quad \frac{\partial r}{\partial y} = \sin\theta$$

et de même avec la seconde ligne

$$\frac{\partial\theta}{\partial x} = -\frac{\sin\theta}{r} \quad \text{et} \quad \frac{\partial\theta}{\partial y} = \frac{\cos\theta}{r}$$

ce qui peut s'écrire aussi

$$d\theta = -\frac{r\sin\theta}{r^2}dx + \frac{r\cos\theta}{r^2}dy = \frac{1}{x^2+y^2}(-ydx + xdy)$$

Comme inverser une matrice revient à résoudre un système linéaire, les calculs précédents peuvent se mener de manière formelle. C'est encore un des avantages de la notation différentielle :

$$\begin{cases} dx = \cos\theta.dr - r\sin\theta.d\theta \\ dy = \sin\theta.dr + r\cos\theta.d\theta \end{cases} \Rightarrow \begin{cases} dr = \cos\theta.dx + \sin\theta.dy \\ d\theta = -\dfrac{\sin\theta}{r}dx + \dfrac{\cos\theta}{r}dy \end{cases}$$

REMARQUE 18-2.17 Le calcul précédent montre bien que, dans le symbole $\dfrac{\partial x}{\partial r}$ qui représente la dérivée partielle d'une fonction par rapport à une variable, il serait ridicule de séparer ∂r et ∂x, et de traiter formellement cette dérivée partielle comme un quotient ! En effet, lorsqu'on considère l'application Φ^{-1},

$$\frac{\partial r}{\partial x} \text{ n'est pas l'inverse de } \frac{\partial x}{\partial r}$$

Affirmer le contraire, c'est finalement penser que, pour inverser une matrice, on remplace chaque coefficient par son inverse ...

EXERCICE 18-2.18 Si α est un réel fixé, trouver toutes les fonctions f de classe C^1 de $]0,+\infty[\times]0,+\infty[\to \mathbb{R}$ vérifiant

$$\forall x > 0, y > 0 \quad x\frac{\partial f}{\partial x}(x,y) + y\frac{\partial f}{\partial y}(x,y) - \alpha f(x,y) = (x^2+y^2)$$

18-2.5 Fonctions de classe C^k

18-2.5.1 Dérivée partielle d'ordre supérieur à 1. Théorème de Schwarz

DÉFINITION 18-2.19 *Soit f définie sur un ouvert U de $\mathbb{E}_n$ à valeurs dans $\mathbb{F}_p$. Si $B = (e_i)_{1\leqslant i\leqslant n}$ est une base fixée de $\mathbb{E}_n$, on suppose que la fonction $D_i f = \dfrac{\partial f}{\partial x_i}$ existe au voisinage d'un point $a \in U$. Si cette fonction possède en a une dérivée partielle par rapport à la $j^{\text{ème}}$ variable, on dit que f possède une dérivée partielle d'ordre 2 par rapport à la $i^{\text{ème}}$ puis la $j^{\text{ème}}$ variable, que l'on note*

$$D_{ji}f(a) = D_j(D_i f)(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

Lorsque $i = j$, on notera aussi

$$D_{ii}f(a) = \frac{\partial^2 f}{\partial x_i^2}(a)$$

On définira de même des dérivées partielles d'ordre 3, 4 etc... Avec un espace à n dimensions où une base est fixée, on peut envisager *a priori* en un point a donné, lorsqu'elles existent, n dérivées partielles d'ordre 1, n^2 dérivées partielles d'ordre 2, n^3 dérivées partielles d'ordre 3 etc... Le théorème qui suit permet souvent de permuter l'ordre des dérivations, et donc par exemple de n'envisager en un point a que le calcul de $\dfrac{n(n+1)}{2}$ dérivées partielles d'ordre 2.

THÉORÈME 18-2.20 Soit $f : U \to \mathbb{F}_p$ admettant au voisinage d'un point a des dérivées partielles secondes $D_{ij}f$ et $D_{ji}f$. Si ces dérivées partielles sont continues en a, alors

$$D_{ij}f(a) = D_{ji}f(a)$$

Démonstration : En travaillant composante par composante dans une base de $\mathbb{F}_p$, on peut supposer f à valeurs réelles. De plus, en considérant la fonction (définie au voisinage de l'origine dans $\mathbb{R}^2$)

$$(x,y) \mapsto f(a + xe_i + ye_j)$$

on se ramène à travailler avec une fonction définie au voisinage de l'origine dans $\mathbb{R}^2$ (que nous noterons encore f). Il existe un réel $\alpha > 0$ tel que les dérivées partielles $D_1 f$, $D_2 f$, $D_{21}f$ et $D_{12}f$ existent sur $[0,\alpha]\times[0,\alpha]$. Pour $0 < h < \alpha$ et $0 < k < \alpha$, définissons

$$\Delta(h,k) = [f(h,k) - f(h,0)] - [f(0,k) - f(0,0)]$$

Cette quantité est accroissement sur le segment $[0, h]$ de la fonction dérivable

$$\varphi(t) = f(t,k) - f(t,0)$$

avec, par définition de la dérivée particlle

$$\varphi'(t) = D_1 f(t,k) - D_1 f(t,0)$$

D'après le théorème des accroissements finis, il existe un réel de $[0,1]$ dépendant de h et k et que nous noterons $\theta_1(h,k)$ tel que

$$\Delta(h,k) = h\varphi'(\theta_1(h,k)\,h) = h\,[D_1 f(h\theta_1,k) - D_1 f(h\theta_1,0)]$$

La quantité entre crochets est accroissement sur $[0,k]$ de la fonction dérivable $\psi(t) = D_1 f(h\theta_1, t)$, avec, par définition de la dérivée partielle d'ordre 2, $\psi'(t) = D_{21} f(h\theta_1, t)$. Ici encore, par application de la formule des accroissements finis, il existe $\theta_2(h,k) \in [0,1]$ vérifiant

$$\Delta(h,k) = hkD_{21} f(h\theta_1, k\theta_2)$$

Mais en écrivant

$$\Delta(h,k) = [f(h,k) - f(0,k)] - [f(h,0) - f(0,0)]$$

on permute les rôles joués par les deux variables, et on obtiendrait de même l'existence de deux réels $\theta_1'(h,k)$ et $\theta_2'(h,k) \in [0,1]$ tels que

$$\Delta(h,k) = hkD_{12} f(h\theta_2', k\theta_1')$$

Comme $D_{12}f$ et $D_{21}f$ sont supposées être continues en $(0,0)$, on en déduit aisément

$$\lim_{h\to 0^+} \frac{\Delta(h,h)}{h^2} = D_{12}f(0,0) = D_{21}f(0,0) \quad \blacksquare$$

18-2.5.2 Fonctions de classe C^k

DÉFINITION 18-2.21 *Soit $f : U \to \mathbb{F}_p$. On dit que f est une fonction de classe C^k dans l'ouvert U si et seulement s'il existe un base $\mathcal{B}$ de $\mathbb{E}_n$ dans laquelle la fonction f possède en tout point de U des dérivées partielles d'ordre inférieur ou égal à k par rapport à toutes les variables (n dérivées d'ordre $1, n^2$ dérivées d'ordre $2,\ldots, n^k$dérivées d'ordre k), ces dérivées partielles étant toutes continues sur U.*

Montrons d'abord, par récurrence sur k, que cette propriété est intrinsèque, c'est-à-dire que si elle est vérifiée par une fonction dans une base de $\mathbb{E}_n$, elle l'est dans toutes les bases de $\mathbb{E}_n$. Notons, pour une fonction $g : U \to \mathbb{F}_p$, la propriété

$P(g,k,\mathcal{B})$: la fonction g vérifie la définition précédente au rang k dans la base $\mathcal{B}$

et montrons, par récurrence sur k, la propriété

$$P(k) : \quad \forall g : U \to \mathbb{F}_p \qquad [(\exists \mathcal{B} \text{ base} \quad P(g,k,\mathcal{B})) \Rightarrow (\forall \mathcal{B}' \text{ base} \quad P(g,k,\mathcal{B}'))]$$

Pour $k = 1$, cela a déjà été vu. Supposons $k \geqslant 2$ et la propriété $P(k-1)$ vérifiée. Donnons nous une fonction f vérifiant $P(f, k, \mathcal{B})$. Soit $\mathcal{B}'$ une autre base de $\mathbb{E}_n$. Notons $D_i f$ et $D'_i f$ les dérivées partielles de f dans $\mathcal{B}$ et $\mathcal{B}'$ (elles existent, puisque f est au moins $\mathcal{C}^1$). Par hypothèse, on a

$$\forall i = 1 \ldots n \quad P(D_i f, k-1, \mathcal{B})$$

et donc, par hypothèse de récurrence

$$\forall i = 1 \ldots n \quad P(D_i f, k-1, \mathcal{B}')$$

Si on a $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ et $\mathcal{B}' = (e'_i)_{1 \leqslant i \leqslant n}$, avec $e'_i = \sum_{j=1}^{n} \alpha_{ij} e_j$, nous avons

$$\forall a \in U \quad D'_i f(a) = df_a(e'_i) = \sum_{j=1}^{n} \alpha_{ij} df_a(e_j) = \sum_{j=1}^{n} \alpha_{ij} D_j f(a)$$

ce qui montre que les fonctions $D'_i f$ sont combinaisons linéaires des $D_j f$. Comme il est clair que

$$\{g : U \to \mathbb{F}_p \mid P(g, k-1, \mathcal{B}')\} \text{ est un sous-espace vectoriel de } \mathcal{C}^1(U, \mathbb{F}_p)$$

on obtient

$$\forall i = 1 \ldots n \quad P(D'_i f, k-1, \mathcal{B}')$$

ce qui traduit exactement $P(f, k, \mathcal{B}')$.

$\mathcal{C}^k(U, \mathbb{F}_p)$ désignera l'ensemble des applications de classe $\mathcal{C}^k$ de U dans $\mathbb{F}_p$, et comme d'habitude

$$\mathcal{C}^\infty(U, \mathbb{F}_p) = \bigcap_{k \in \mathbb{N}} \mathcal{C}^k(U, \mathbb{F}_p)$$

Comme les calculs de dérivées partielles font intervenir, à chaque étape de dérivation, des calculs sur des fonctions d'une variable réelle (application partielle de la fonction étudiée), on a évidemment

PROPOSITION 18-2.22 **$\mathcal{C}^k(U, \mathbb{F}_p)$ est un sous-espace vectoriel de $\mathcal{C}^0(U, \mathbb{F}_p)$. De même, grâce à la formule de Leibniz, $\mathcal{C}^k(U, \mathbb{K})$ est une sous-algèbre de $\mathcal{C}^0(U, \mathbb{K})$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Il en est de même de $\mathcal{C}^\infty(U, \mathbb{K})$.**

Plus généralement, si $\varphi : U \to \mathbb{R}$ et $f : U \to \mathbb{F}_p$ sont deux applications de classe $\mathcal{C}^k$, le produit φf est aussi de classe $\mathcal{C}^k$.

PROPOSITION 18-2.23 **En fixant une base $\mathcal{B}_1 = (u_i)_{1 \leqslant i \leqslant p}$ de $\mathbb{F}_p$ et en considérant les applications composantes de $f : U \to \mathbb{F}_p$**

$$f = \sum_{i=1}^{p} f_i u_i$$

on a évidemment : f est de classe $\mathcal{C}^k$ sur U si et seulement si toutes les f_i le sont.

PROPOSITION 18-2.24 Toute application polynomiale $\mathbb{R}^n \to \mathbb{R}$ est dans $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$

Par applications successives du théorème de Schwarz, les dérivées partielles d'ordre $\leqslant k$ d'une fonction de classe $\mathcal{C}^k$ sur un ouvert U peuvent être évaluées après permutation de l'ordre de dérivation : si $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ est un multi-indice de longueur

$$p = |\alpha| = \sum_{i=1}^n \alpha_i \leqslant k$$

nous noterons

$$\frac{\partial^p f}{\partial x^\alpha} = \frac{\partial^p f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

une dérivée partielle de f obtenue par α_1 dérivations par rapport à la première variable, α_2 dérivations par rapport à la seconde etc...

PROPOSITION 18-2.25 Si U est un ouvert de $\mathbb{E}_n$ et V un ouvert de $\mathbb{F}_p$, et $f \in \mathcal{C}^k(U, \mathbb{F}_p)$, $g \in \mathcal{C}^k(V, \mathbb{G}_q)$ avec $f(U) \subset V$, alors $g \circ f \in \mathcal{C}^k(U, \mathbb{G}_q)$.

Démonstration : La démonstration se fait par récurrence sur k. Pour $k = 1$, c'est simplement le corollaire **18-2.**11. Supposons $k \geqslant 2$ et le résultat établi pour les fonctions de classe $\mathcal{C}^{k-1}$. Soient deux fonctions $f \in \mathcal{C}^k(U, \mathbb{F}_p)$ et $g \in \mathcal{C}^k(V, \mathbb{G}_q)$. Fixons des bases $\mathcal{B} = (e_j)_{1 \leqslant j \leqslant n}$ et $\mathcal{B}_1 = (u_i)_{1 \leqslant i \leqslant p}$ de $\mathbb{E}_n$ et $\mathbb{F}_p$ et caractérisons f par ses applications composantes

$$f = \sum_{i=1}^p f_i u_i$$

avec $f_i \in \mathcal{C}^k(U, \mathbb{R})$. On a alors, par application du théorème de dérivation d'une fonction composée

$$\forall a \in U \quad \forall j = 1, \dots, n \quad D_j(g \circ f)(a) = \sum_{i=1}^p D_i g(f(a)) . D_j f_i(a)$$

Pour $1 \leqslant i \leqslant p$, l'application $(D_i g) \circ f$ est de classe $\mathcal{C}^{k-1}$ d'après l'hypothèse de récurrence. Comme $D_j f_i$ possède la même propriété, le produit $(D_j f_i)((D_i g) \circ f)$ est également de classe $\mathcal{C}^{k-1}$, et par linéarité, on obtient

$$\forall j = 1, \dots, n \quad D_j(g \circ f) \in \mathcal{C}^{k-1}(U, \mathbb{G}_q)$$

ce qui montre bien que $g \circ f \in \mathcal{C}^k(U, \mathbb{G}_q)$. ■

Comme l'application $x \mapsto \dfrac{1}{x}$ est $\mathcal{C}^\infty$ de $\mathbb{R}^*$ dans lui-même, on en déduit

COROLLAIRE 18-2.26 Si $f \in \mathcal{C}^k(U,\mathbb{R})$ ne s'annule pas sur U, alors

$$\frac{1}{f} \in \mathcal{C}^k(U,\mathbb{R})$$

En particulier, toute fraction rationnelle de n variables réelles est de classe $\mathcal{C}^\infty$ sur son ouvert de définition.

EXERCICE 18-2.27 Montrer que l'application $M \mapsto M^{-1}$ est de classe $\mathcal{C}^\infty$ de $\mathcal{GL}_n(\mathbb{R})$ dans lui-même.

18-3 Applications des différentielles

18-3.1 Inégalité des accroissements finis

18-3.1.1 Théorème des accroissements finis

THÉORÈME 18-3.1 Soit U un ouvert de $\mathbb{E}_n$ et $f \in \mathcal{C}^1(U,\mathbb{F}_p)$. Si a et b sont deux points de U tels que le segment $[a,b]$ soit inclus dans U alors

$$\|f(b)-f(a)\| \leqslant \left(\sup_{x\in[a,b]} \|df_x\|\right).\|b-a\|$$

Démonstration : Dans l'inégalité à montrer, $\|df_x\|$ représente évidemment la norme de l'application linéaire continue $df_x \in \mathcal{L}(\mathbb{E}_n,\mathbb{F}_p)$ (norme subordonnée aux normes choisies sur $\mathbb{E}_n$ et $\mathbb{F}_p$, qui interviennent également dans l'inégalité). Remarquons que, comme l'application $x \mapsto df_x$ est continue sur le compact $[a,b]$, $M = \sup_{x\in[a,b]} \|df_x\|$ existe bien. Comme souvent, pour démontrer des résultats sur les fonctions de plusieurs variables, on se ramène à travailler en dimension 1. Ici, utilisant l'hypothèse $[a,b] \subset U$, posons, pour $t \in [0,1]$,

$$\varphi(t) = f(a+t(b-a))$$

Cette fonction est de classe $\mathcal{C}^1$ sur $[0,1]$, et vérifie

$$\begin{aligned}\forall t \in [0,1] \quad \|\varphi'(t)\| &= \left\|df_{a+t(b-a)}(b-a)\right\| \\ &\leqslant \left\|df_{a+t(b-a)}\right\| \|b-a\| \leqslant M\|b-a\|\end{aligned}$$

L'inégalité des accroissements finis donne alors

$$\|\varphi(1)-\varphi(0)\| \leqslant M\|b-a\|$$

ce qui est exactement l'inégalité à démontrer. ■

REMARQUE 18-3.2 Lorsqu'on travaille dans une base $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ de $\mathbb{E}_n$, et donc avec les dérivées partielles correspondantes, on obtient, pour $b = a+h$ avec $h = \sum_{i=1}^{n} h_i e_i$, et toujours sous l'hypothèse $[a,a+h] \subset U$

$$\varphi'(t) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a+th)$$

et donc, si on connaît des constantes $(M_i)_{1\leqslant i\leqslant n}$ telles que

$$\left\|\frac{\partial f}{\partial x_i}\right\| \leqslant M_i \text{ sur } [a, a+h]$$

l'inégalité des accroissements finis donnera dans ce cas

$$\|f(a+h) - f(a)\| \leqslant \sum_{i=1}^{n} M_i |h_i|$$

REMARQUE 18-3.3 Dans le cas d'une fonction réelle, l'égalité des accroissements finis appliquée à φ donnera

$$\exists\, \theta \in \left]0,1\right[\quad f(a+h) - f(a) = df_{a+\theta h}(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a+\theta h)$$

18-3.1.2 Applications

THÉORÈME 18-3.4 Soit U un ouvert convexe de $\mathbb{E}_n$. Si k est un réel positif et $f : U \to \mathbb{F}_p$ est une fonction de classe $\mathcal{C}^1$,

$$f \textbf{ est } k\textbf{-lipschitzienne sur } U \Leftrightarrow \forall x \in U \quad \|df_x\| \leqslant k$$

Démonstration : Si $\|df_x\| \leqslant k$ en tout point de U, comme U est convexe on a

$$\forall\, a, b \in U \quad [a, b] \subset U$$

et le théorème précédent donne bien

$$\forall\, a, b \in U \quad \|f(b) - f(a)\| \leqslant k\, \|b - a\|$$

Réciproquement, si f est k-lipschitzienne, pour $a \in U$, $x \in \mathbb{E}_n$ et $t \in \mathbb{R}$ assez petit

$$\left\|\frac{f(a+tx) - f(a)}{t}\right\| \leqslant k\, \|x\|$$

En faisant tendre t vers 0, on obtient

$$\forall\, a \in U \quad \forall\, x \in \mathbb{E}_n \quad \|df_a(x)\| \leqslant k\, \|x\|$$

ce qui donne bien $\|df_a\| \leqslant k$ pour tout a dans U. ■

THÉORÈME 18-3.5 Si U est un ouvert connexe de $\mathbb{E}_n$ et $f : U \to \mathbb{F}_p$ est de classe $\mathcal{C}^1$

$$f \textbf{ est constante sur } U \Leftrightarrow \forall x \in U \quad df_x = 0_{\mathcal{L}(\mathbb{E}_n, \mathbb{F}_p)}$$

Démonstration : En travaillant dans une base de $\mathbb{E}_n$, on aura évidemment

$$\forall x \in U \quad df_x = 0_{\mathcal{L}(\mathbb{E}_n,\mathbb{F}_p)} \Leftrightarrow \forall i = 1, \dots, n \quad \forall x \in U \quad \frac{\partial f}{\partial x_i}(x) = 0_{\mathbb{F}_p}$$

Cette condition est évidemment nécessaire pour que f soit constante dans U. Réciproquement, si elle est vérifiée, en appliquant le théorème précédent avec $k = 0$ on s'aperçoit que f est constante dans toute boule ouverte incluse dans U (puisqu'une telle boule est convexe). f est donc localement constante sur l'ouvert connexe U. Elle est donc bien constante sur U (si U n'était pas connexe, f serait constante sur chaque composante connexe de U). ■

REMARQUE 18-3.6 Prenons une fonction de deux variables réelles

$$(x, y) \mapsto F(x, y)$$

définie sur un ouvert U de $\mathbb{R}^2$, admettant en tout point de U une dérivée partielle par rapport à la première variable et vérifiant

$$\frac{\partial F}{\partial x} \equiv 0 \text{ sur } U$$

L'usage est de dire alors

"F ne dépend que de y"

et d'écrire

$$F(x, y) = \varphi(y)$$

Ce genre de raisonnement a besoin d'être précisé : que signifie "F ne dépend que de y" ? Quelle propriété doit vérifier la "fonction" φ ?

Soit V la projection de U sur l'axe des y

$$V = \{y \in \mathbb{R} \,|\, \exists x \in \mathbb{R} \quad (x, y) \in U\}$$

Il est facile de voir que V est un ouvert de $\mathbb{R}$. C'est bien évidemment sur cet ouvert que sera définie, si elle existe, la fonction φ. Cette dernière existera certainement si l'ouvert U est tel que[7]

$$\forall y \in V \quad U_y = \{x \in \mathbb{R} \,|\, (x, y) \in U\} \text{ est un intervalle (ouvert) de } \mathbb{R}$$

En effet, si c'est le cas, pour $y \in V$

$$F(\bullet, y) : U_y \to \mathbb{F}_p$$

est une fonction dérivable sur un intervalle de $\mathbb{R}$, avec une dérivée identiquement nulle. Cette fonction est donc constante sur cet intervalle,, et on peut noter $\varphi(y)$ la valeur de cette fonction. On définit alors une fonction $\varphi : V \to \mathbb{F}_p$ telle que

$$\forall (x, y) \in U \quad F(x, y) = \varphi(y)$$

[7] U_y est la "coupe" de l'ouvert U en y. C'est un ouvert de $\mathbb{R}$. Lorsque U est un pavé ouvert $]a, b[\times]c, d[$, on a $U_y =]a, b[$ pour $y \in V =]c, d[$.

La régularité de F est alors celle de φ : si F est $\mathcal{C}^k$ sur U, il en est de même de φ sur V.

Par contre, si U ne possède pas la propriété évoquée plus haut, le raisonnement précédent n'est valable que localement (par exemple dans tout pavé ouvert inclus dans U). Prendre par exemple

$$U =]-2,-1[\times]-1,1[\cup]-2,2[\times]-1,0[\cup]1,2[\times]-1,1[$$

(faire un dessin !) et $F : U \to \mathbb{R}$ définie par

$$F(x,y) = y^4 \text{ si } x > 0 \text{ et } y > 0 \qquad F(x,y) = y^3 \text{ sinon}$$

$F \in \mathcal{C}^1(U,\mathbb{R})$ et vérifie $\dfrac{\partial F}{\partial x} \equiv 0$. Localement, cette fonction ne dépend que de y, mais c'est faux globalement, puisque le signe de x intervient dans sa définition.

18-3.2 Développement limité d'ordre 2 pour une fonction numérique de classe $\mathcal{C}^2$

Dans cette section, nous ne considérerons, pour simplifier, que des fonctions à valeurs réelles. On pourrait généraliser certains résultats pour des fonctions à valeurs dans un espace de dimension finie, le plus simple étant alors de travailler avec les fonctions composantes dans une base de l'espace d'arrivée.

18-3.2.1 Notion de développement limité d'ordre 2

DÉFINITION 18-3.7 *Soit $f : U \to \mathbb{R}$ et a un point de U. On dit que f admet un développement limité d'ordre 2 en a si et seulement s'il existe une forme linéaire $l \in \mathbb{E}_n^*$ et une forme quadratique $\Phi \in Q(\mathbb{E}_n)$ telles que l'on ait*

$$f(a+h) = f(a) + l(h) + \Phi(h) + \|h\|^2 \varepsilon(h) \quad \textit{avec} \quad \lim_{h \to 0_{\mathbb{E}_n}} \varepsilon(h) = 0$$

Dans une base donnée de $\mathbb{E}_n$, $l(h)$ s'exprime comme polynôme homogène de degré 1 des coordonnées de h, alors que $\Phi(h)$ est polynôme homogène de degré 2.

Qui peut le plus peut le moins : si un tel développement existe, f est différentiable en a et $l = df_a$. En effet, Φ étant continue sur $\mathbb{E}_n$ est bornée sur la sphère unité et, par homogénéité, il existe une constante $M \geqslant 0$ telle que

$$\forall h \in \mathbb{E}_n \quad |\Phi(h)| \leqslant M \|h\|^2$$

et par conséquent $\Phi(h) + \|h\|^2 \varepsilon(h) \underset{h \to 0_{\mathbb{E}_n}}{=} o(\|h\|)$.

PROPOSITION 18-3.8 Si $f : U \to \mathbb{R}$ possède un développement limité d'ordre 2 en $a \in U$, ce développement limité est unique.

Démonstration : Comme pour l'unicité de la différentielle, on étudie f au voisinage de a dans la direction d'un vecteur $x \in \mathbb{E}_n$ arbitraire :

$$f(a+tx) \underset{t \to 0}{=} f(a) + t\,df_a(x) + t^2\Phi(x) + t^2\|x\|^2 \varepsilon(tx)$$

qui donne un développement limité d'une fonction numérique définie au voisinage de 0. Pour $x \in \mathbb{E}_n$, $\Phi(x)$ est donc unique, ce qui prouve l'unicité de Φ. ■

On pourrait généraliser et introduire la notion de développement limité d'ordre supérieur. Par exemple, à l'ordre 3, on demanderait l'existence d'une fonction P_3 des coordonnées de h dans une[8] base de $\mathbf{E}_n$ polynomiale homogène de degré 3 telle que

$$f(a+h) = f(a) + l(h) + \Phi(h) + P_3(h) + \|h\|^3 \varepsilon(h)$$

Ces définitions sont cohérentes : par exemple un DL d'ordre 3 est plus précis qu'un DL d'ordre 2. En effet, un polynôme homogène de degré k

$$P_k : \mathbb{R}^n \to \mathbb{R}$$

est borné sur la sphère unité de $(\mathbb{R}^n, \| \|_\infty)$ (par continuité et compacité) et vérifie donc une inégalité du type

$$\forall x \in \mathbb{R}^n \quad |P_k(x)| \leqslant M \|x\|_\infty^k \text{ et donc } P_k(h) \underset{h \to 0_{\mathbb{E}_n}}{=} o\left(\|h\|^{k-1}\right)$$

Pour l'essentiel, nous nous limiterons dans la suite aux développements limités d'ordre 2.

18-3.2.2 Formule de Taylor-Young

THÉORÈME 18-3.9 Soit $f : U \to \mathbb{R}$ de classe C^2. f possède en tout point a de U un développement limité d'ordre 2. En travaillant dans une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ avec $h = \sum_{i=1}^{n} h_i e_i$, ce développement est donné par la formule de Taylor-Young :

$$f(a+h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2} \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o\left(\|h\|^2\right)$$

Démonstration : Il existe un réel $\alpha > 0$ tel que

$$\|h\| \leqslant \alpha \Rightarrow a + h \in U$$

L'idée est encore ici, pour $\|h\| \leqslant \alpha$, de considérer la fonction

$$\varphi(t) = f(a + th)$$

pour $t \in [0,1]$. φ est de classe C^2 comme composée de fonctions de classe C^2. On a

$$\forall t \in [0,1] \quad \varphi'(t) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a + th)$$

[8]Les formules de changement de base montrent alors que, dans toute base de $\mathbb{E}_n$, $P_3(h)$ s'exprime comme polynôme homogène de degré 3 des coordonnées de h.

et, en appliquant à nouveau la formule de dérivation d'une fonction composée

$$\forall t \in [0,1] \quad \varphi''(t) = \sum_{j=1}^{n} h_j \left(\sum_{i=1}^{n} h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(a+th) \right)$$

On écrit la formule de Taylor avec reste intégrale pour φ sur $[0,1]$ sous la forme

$$f(a+h) = \varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 (1-t)\varphi''(t)\,dt$$
$$= \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(0) + R(h)$$

avec

$$R(h) = \int_0^1 (1-t)\left[\varphi''(t) - \varphi''(0)\right] dt$$
$$= \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} h_i h_j \int_0^1 (1-t)\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a+th) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right] dt$$

Compte tenu des valeurs des valeurs de $\varphi(0)$, $\varphi'(0)$ et $\varphi''(0)$, la formule sera démontrée si nous prouvons que $R(h) \underset{h \to 0_{E_n}}{=} o\left(\|h\|^2\right)$. Ceci se ramène à prouver que, pour i et j quelconques

$$\lim_{h \to 0_{E_n}} \int_0^1 (1-t)\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a+th) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right] dt = 0$$

ce qui est conséquence simple du théorème de continuité d'une intégrale dépendant d'un paramètre. ■

Par application du théorème de Schwarz, les termes rectangles peuvent être regroupés. Par exemple, pour une fonction de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$, nous aurons

$$f(x_0+h, y_0+k) = f(x_0,y_0) + h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0) +$$
$$\frac{1}{2}\left(h^2\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2hk\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right) + \left(h^2+k^2\right)\varepsilon(h,k)$$

avec $\lim\limits_{(h,k)\to(0,0)} \varepsilon(h,k) = 0$.

La matrice (dans la base $\mathcal{B}$) de la forme quadratique qui apparaît dans le développement limité de f en a est la matrice symétrique

$$\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

DÉFINITION 18-3.10 *Si f est une fonction numérique de classe C^2 sur un ouvert U de $\mathbb{E}_n$, ramené à une base $\mathcal{B}$, on appelle matrice Hessienne de f en a (en précisant "dans la base $\mathcal{B}$" si besoin est) la matrice symétique d'ordre n*

$$H(f)(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$$

En notant X la colonne des coordonnées de h dans $\mathcal{B}$, le développement limité de f en a s'écrit alors

$$f(a+h) = f(a) + L(a)X + \frac{1}{2}{}^t X\,[H(f)(a)]\,X + o\left(\|X\|^2\right)$$

où $L(a)$ est la matrice (dans $\mathcal{B}$) de la différentielle de f en a :

$$L(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)$$

Effet d'un changement de base :

Si P est matrice de passage de la base $\mathcal{B}$ à une autre base $\mathcal{B}_1$ de $\mathbb{E}_n$, et si on note $H^1(f)(a)$ la matrice Hessienne de f dans $\mathcal{B}_1$, comme ces deux matrices représentent la même forme quadratique dans des bases distinctes, on a

$$H^1(f)(a) = {}^t P\,[H(f)(a)]\,P$$

conformément à ce qui a été vu dans le chapitre sur les formes quadratiques.

EXERCICE 18-3.11 Enoncer et démontrer un résultat sur le développement limité à l'ordre k d'une fonction de classe C^k avec $k > 2$.

18-3.2.3 Application à la recherche d'extrema d'une fonction numérique

DÉFINITION 18-3.12 *Soit A une partie d'un espace normé et $f : A \to \mathbb{R}$ une fonction numérique. Si $x_0 \in A$, on dit que f présente un maximum local en x_0 si l'on peut trouver un voisinage[9] V de x_0 tel que*

$$\forall x \in V \cap A \quad f(x) \leqslant f(x_0)$$

Par abus de langage, on dira aussi que x_0 est un maximum local de f. On dit que ce maximum est strict si on peut trouver un tel V vérifiant

$$\forall x \in V \cap A - \{x_0\} \quad f(x) < f(x_0)$$

Enfin, on parlera de maximum absolu ou global si ces inégalités sont valables respectivement sur A et $A - \{x_0\}$. En changeant le sens des inégalités, on définirait de même un minimum (local, strict, global). Un extremum de f est un maximum ou un minimum.

[9] Cette notion n'a évidemment d'intérêt que si x_0 n'est pas un point isolé de A, c'est à dire si tout voisinage de x_0 rencontre $A - \{x_0\}$.

Rappelons que la compacité est un argument souvent utilisé pour justifier l'existence d'un extremum global. Une fonction numérique **continue sur un compact** est bornée et atteint ses bornes.

Comme dans le cas des fonctions d'une variable réelle, le calcul différentiel donne des conditions nécessaires pour qu'un point x_0 corresponde à un extremum local d'une fonction f :

THÉORÈME 18-3.13 Soient $f : A \to \mathbb{R}$ et $x_0 \in A$. Si A est un voisinage de x_0, si f présente un extremum local en x_0 et est différentiable en x_0 alors $df_{x_0} = 0_{\mathbb{E}_n^*}$. En travaillant dans une base de $\mathbb{E}_n$, on a en particulier

$$\forall i \in \{1 \dots n\} \quad \frac{\partial f}{\partial x_i}(x_0) = 0$$

> *Démonstration* : La fonction d'une variable réelle définie au voisinage de 0 par $\varphi(t) = f(x_0 + th)$ (où h est un vecteur arbitraire de $\mathbb{E}_n$) est dérivable en 0, où elle présente un extremum local. On a donc $\varphi'(0) = 0$, ce qui donne
>
> $$\forall h \in \mathbb{E}_n \quad df_{x_0}(h) = 0 \quad \blacksquare$$

Notons bien l'hypothèse importante : A est un voisinage de x_0. Par exemple dans $\mathbb{R}^3$ euclidien canonique, la fonction f définie par $f(x, y, z) = x + 2y + 3z$ est de classe C^∞. Sa restriction à la sphère unité S atteint un maximum et un minimum global (par compacité) mais la différentielle de f ne s'annule en aucun point de S.

DÉFINITION 18-3.14 *Si f est une fonction numérique de classe C^1 sur un ouvert U de $\mathbb{E}_n$, un point $x_0 \in U$ est point critique de f si et seulement si la différentielle de f est nulle en x_0.*

Les extrema locaux éventuels de f sont alors à rechercher parmi les points critiques de f. Remarquons que la condition $df_{x_0} = 0_{\mathbb{E}_n^*}$ n'est pas suffisante pour assurer que x_0 soit extremum local de f. Ceci est déjà bien connu en dimension 1, avec l'exemple $x \to x^3$ en 0.

Le théorème qui précède pourra être appliqué de la manière suivante :

COROLLAIRE 18-3.15 Si K est un compact de $\mathbb{E}_n$ et $f : K \to \mathbb{R}$ une fonction continue. Si f est de classe C^1 sur $\mathring{K}$, intérieur de K (ce qui est certainement le cas si f est prolongeable en une fonction C^1 sur un ouvert contenant K), les points de K où f atteint ses extrema globaux sont à rechercher parmi les points critiques de $f|_{\mathring{K}}$ et parmi les points-frontière de K.

Dans la suite, nous supposerons f définie sur un ouvert U et de classe C^2. L'utilisation d'un développement limité au voisinage d'un point critique donnera une condition nécessaire et une condition suffisante d'existence d'un extremum local :

PROPOSITION 18-3.16 Soit $f : U \to \mathbb{R}$ de classe C^2, et $x_0 \in U$. Si x_0 est un maximum local de f, la forme quadratique associée à la matrice Hessienne de f en x_0 est négative. Si x_0 est minimum local de f, cette forme quadratique est positive.

Démonstration : Supposons que x_0 soit un maximum de f et notons Φ la forme quadratique associée à $H(f)(x_0)$. On a donc

$$f(x_0 + h) = f(x_0) + \frac{1}{2}\Phi(h) + \|h\|^2 \varepsilon(h)$$

avec $\lim_{h \to 0_{\mathbb{E}_n}} \varepsilon(h) = 0$, puisque x_0 maximum est nécessairement point critique de f. Supposons que Φ ne soit pas négative. Il existe alors $x \in \mathbb{E}_n$ avec $\Phi(x) > 0$. On a alors, pour t réel tendant vers 0,

$$f(x_0 + tx) = f(x_0) + \frac{t^2}{2}\Phi(x) + o(t^2)$$

et donc

$$f(x_0 + tx) - f(x_0) \underset{t \to 0}{\sim} \frac{t^2}{2}\Phi(x) > 0$$

ce qui est contradictoire avec le fait que x_0 soit un maximum local de f. Le cas x_0 minimum s'étudie de façon analogue. ■

Il résulte de cette étude qu'un point critique x_0 d'une fonction f de classe C^2 pour lequel la matrice symétrique $H(f)(x_0)$ n'est ni positive ni négative ne peut être extremum local de f.

PROPOSITION 18-3.17 Soit x_0 un point critique d'une fonction $f : U \to \mathbb{R}$ de classe C^2. Si la forme quadratique associée à la matrice Hessienne de f en x_0 est définie négative, alors x_0 est un maximum local de f. Si cette forme quadratique est définie positive, x_0 est un minimum local de f.

Démonstration : Supposons cette forme quadratique Φ définie positive. On a

$$f(x_0 + h) = f(x_0) + \frac{1}{2}\Phi(h) + \|h\|^2 \varepsilon(h)$$

avec $\lim_{h \to 0_{\mathbb{E}_n}} \varepsilon(h) = 0$. Comme Φ est une forme quadratique définie positive, sa racine carrée définit une norme sur $\mathbb{E}_n$, qui est équivalente à $\| \, \|$. Il existe donc une constante $M > 0$ avec

$$\forall x \in \mathbb{E}_n \quad \Phi(x) \geqslant M \|x\|^2$$

et en conséquence

$$f(x_0 + h) - f(x_0) \geqslant \left(\frac{M}{2} + \varepsilon(h)\right) \|h\|^2$$

Comme ε tend vers 0 à l'origine, on peut trouver un réel $\alpha > 0$ tel que

$$\|h\| \leqslant \alpha \Rightarrow x_0 + h \in U \text{ et } |\varepsilon(h)| \leqslant \frac{M}{4}$$

On a alors, pour $0 < \|h\| \leqslant \alpha$

$$f(x_0 + h) - f(x_0) \geqslant \frac{M}{4} \|h\|^2 > 0$$

ce qui prouve bien que x_0 est un minimum local de f. Le cas où Φ est définie négative est analogue. ■

Les deux propositions précédentes montrent que dans le cas où x_0 est point critique de f avec $H(f)(x_0)$ inversible (ce qui veut dire que Φ est non dégénérée, on dit aussi que x_0 est point critique non dégénéré), on a trois possibilités : si Φ a pour signature $(n, 0)$, x_0 est un minimum local ; si la signature de Φ est $(0, n)$, on a un maximum local. Enfin, si la signature est (p, q) avec $p + q = n$ et $pq \neq 0$, x_0 n'est pas un extremum local. En particulier, dans le cas de la dimension 2 :

COROLLAIRE 18-3.18 Soit f de classe C^2 sur un ouvert U de $\mathbb{R}^2$, à valeurs réelles. Si $m_0 = (x_0, y_0) \in U$, notons

$$H(f)(m_0) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} (m_0) = \begin{pmatrix} r_0 & s_0 \\ s_0 & t_0 \end{pmatrix}$$

(notations dites de Monge).

Si m_0 est un point critique de f, alors

$r_0 t_0 - s_0^2 > 0 \Rightarrow m_0$ **est un extremum local de f**
$r_0 t_0 - s_0^2 < 0 \Rightarrow m_0$ **n'est pas un extremum local de f**
$r_0 t_0 - s_0^2 = 0$ **nécessite une étude particulière.**

Démonstration : La signature de Φ est donnée par le signe des valeurs propres de la matrice $H(f)(m_0)$. Le produit de ces valeurs propres est $\det H(f)(m_0) = r_0 t_0 - s_0^2$. Si cette quantité est strictement positive, les deux valeurs propres sont de même signe, et Φ est définie. Le point m_0 est en conséquence un extremum local de f (un maximum si Φ est négative, donc si $r_0 < 0$, un minimum si $r_0 > 0$). Lorsque $r_0 t_0 - s_0^2 < 0$, la signature de Φ est $(1, 1)$, et m_0 n'est pas un extremum local de f. Lorsque $r_0 t_0 - s_0^2 = 0$, la signature de Φ est $(1, 0)$ ou $(0, 1)$, et on ne peut conclure sans une étude plus approfondie de f. ■

EXEMPLE 18-3.19 Etudier le point critique $(0, 0)$ pour les fonctions définies sur $\mathbb{R}^2$ par

$$f(x, y) = x^2 - 4xy + 4y^2 + x^3 + y^3 \text{ et } g(x, y) = x^2 - 4xy + 4y^2 + x^4 + y^4$$

(on remarquera qu'ici, le calcul des dérivées partielles est inutile : f et g sont déjà écrites comme combinaisons linéaires de polynômes homogènes).

EXERCICE 18-3.20 Les fonctions définies sur $\mathbb{R}^2$ et $\mathbb{R}^3$ par

$$f(x,y) = xe^y + ye^x \text{ et } g(x,y,z) = (x+z^2)\, e^{x(y^2+z^2+1)}$$

possèdent-elles des extrema locaux ?

EXERCICE 18-3.21 Montrer que, lorsqu'on la restreint à une droite quelconque passant par l'origine, la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = y^2 - 3x^2y + 2x^4$ présente un minimum en $(0,0)$. Ce point est-il minimum local de f ?

18-3.2.4 Position d'une surface par rapport à son plan tangent

Soit f de classe $\mathcal{C}^2$ définie sur un ouvert U de $\mathbb{R}^2$ et (S) la surface d'équation $z = f(x,y)$ dans un espace affine de dimension 3 ramené à un repère $\left(O,\vec{\imath},\vec{\jmath},\vec{k}\right)$ (voir la section **18-2.3.2**). Soit $M_0(x_0,y_0,z_0 = f(x_0,y_0))$ un point de (S). En notant

$$p_0 = \frac{\partial f}{\partial x}(x_0,y_0) \text{ et } q_0 = \frac{\partial f}{\partial y}(x_0,y_0)$$

on sait que le plan T_0 tangent à (S) en M_0 a pour équation

$$(Z - z_0) = p_0(X - x_0) + q_0(Y - y_0)$$

Si $M(x,y,z)$ est un point quelconque de l'espace, le signe de la quantité

$$\Delta = (z - z_0) - [p_0(x - x_0) + q_0(y - y_0)]$$

permet de placer M par rapport à T_0 ("au-dessus" de T_0 si $\Delta > 0$, "en-dessous" si $\Delta < 0$). Nous étudions ici la position d'un point de (S) "voisin" de M_0, de coordonnées

$$M(x_0 + h, y_0 + k, f(x_0 + h, y_0 + k))$$

donc le signe de

$$\delta(h,k) = f(x_0 + h, y_0 + k) - f(x_0,y_0) - p_0h - q_0k$$

Comme f est supposée $\mathcal{C}^2$, on a, avec les notations de la section précédente

$$\delta(h,k) = \frac{1}{2}\left(h^2r_0 + 2hks_0 + k^2t_0\right) + \left(h^2 + k^2\right)\varepsilon(h,k)$$

et, comme pour les extrema de fonctions de deux variables (l'étude est analogue), on aura

PROPOSITION 18-3.22 Si $r_0t_0 - s_0^2 > 0$, la surface (S) reste, au voisinage de M_0, dans un des demi-espaces délimités par T_0. Le point M_0 est dit elliptique, ou d'aspect "ballon".

Si $r_0t_0 - s_0^2 < 0$, la surface (S) traverse au voisinage de M_0 le plan T_0 : si le point $m_0(x_0,y_0)$ est "légèrement" déplacé dans U dans la direction d'un vecteur (h_0,k_0) avec $\Phi(h_0,k_0) > 0$ selon

$$m_t(x,y) = m_0 + t(h_0,k_0)$$

le point correspondant sur (S) est situé au dessus du plan tangent. Il est situé en dessous si $\Phi(h_0, k_0) < 0$. Le point M_0 est alors appelé "point-selle" ou "point-col" de (S). On dit que M_0 est un point hyperbolique[10]**.**

Si $r_0t_0 - s_0^2 = 0$, une étude particulière est nécessaire pour placer la surface par rapport à son plan tangent.

EXERCICE 18-3.23 Montrer que la surface (hyperboloïde à deux nappes) d'équation cartésienne, dans un repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

ne contient aucun segment de droite (non réduit à un point).

18-3.3 Notion de C^k-difféomorphisme ($k \geqslant 1$)

Nous précisons ici les résultats de la section **18-2.4**.

18-3.3.1 Définition

DÉFINITION 18-3.24 *Soient U et V deux ouverts de deux espaces normés $\mathbb{E}_n$ et $\mathbb{F}_n$ de même dimension n (en général, ce sera $\mathbb{E}_n = \mathbb{F}_n = \mathbb{R}^n$, ce qu'on peut toujours supposer en fixant des bases) et k un entier $\geqslant 1$. Une application $f : U \to V$ est un C^k-difféomorphisme de U dans V si et seulement si f est bijective et si f et f^{-1} sont toutes deux de classe C^k.*

Dans le cas $\mathbb{E}_n = \mathbb{F}_n = \mathbb{R}^n$, une telle application peut être considérée comme un changement de variable de classe C^k :

$$f : U \to V \quad (x_1, \ldots, x_n) \mapsto (y_1 = f_1(x_1, \ldots, x_n), \ldots, y_n = f_n(x_1, \ldots, x_n))$$

le changement de variable réciproque étant également C^k. Un tel changement de variable réalise un isomorphisme d'espaces vectoriels

$$C^k(V, \mathbb{G}_q) \to C^k(U, \mathbb{G}_q) \qquad \Phi \mapsto \Psi = \Phi \circ f$$

On peut ainsi transporter une équation dans $C^k(V, \mathbb{G}_q)$ en une équation dans $C^k(U, \mathbb{G}_q)$, et transporter ensuite les solutions par $\Psi \mapsto \Phi = \Psi \circ f^{-1}$.

Le corollaire **18-2**.15 peut s'écrire, dans le cas des fonctions C^k

> "Si $f : U \to V$ est un homéomorphisme de classe C^k dont la différentielle est inversible en tout point de U, alors f est un C^k-difféomorphisme de U dans V."

[10] On peut montrer (à l'aide du théorème des fonctions implicites) que, au voisinage de M_0, l'intersection $(S) \cap T_0$ est réunion des supports de deux arcs paramétrés C^2 sécants en M_0 dont les tangentes en ce point sont dirigées par les vecteurs

$$\vec{t_1} = h_1\vec{\imath} + k_1\vec{\jmath} + (p_0h_1 + q_0k_1)\vec{k} \text{ et } \vec{t_2} = h_2\vec{\imath} + k_2\vec{\jmath} + (p_0h_2 + q_0k_2)\vec{k}$$

où (h_1, k_1) et (h_2, k_2) sont deux vecteurs indépendants qui annulent la forme quadratique Φ, de signature $(1, 1)$. Ces arcs paramétrés découpent sur (S) des "secteurs angulaires" situés alternativement de part et d'autre du plan tangent T_0.

La démonstration est la même, et fait intervenir le caractère $\mathcal{C}^\infty$ de l'application $M \mapsto M^{-1}$ de $\mathcal{GL}_n(\mathbb{R})$ dans lui-même. Par exemple, le passage en coordonnées polaires Φ étudié à l'exemple **18-2**.16 est un $\mathcal{C}^\infty$-difféomorphisme.

Le théorème d'inversion globale qui suit montre que, dans l'énoncé qui précède, la continuité de f^{-1} est en fait conséquence des autres hypothèses faites sur f. Ce résultat est, bien entendu, à rapprocher du théorème **8-1**.9 et des énoncés de la section **8-4**.22.

18-3.3.2 Théorème d'inversion globale

THÉORÈME 18-3.25 Soit $f : U \to \mathbb{F}_n$ de classe $\mathcal{C}^k$ une application injective définie sur un ouvert U de $\mathbb{E}_n$. Les propriétés suivantes sont équivalentes :

i) $V = f(U)$ est un ouvert de $\mathbb{F}_n$ et f est un $\mathcal{C}^k$-difféomorphisme de U vers V.

ii) En tout point a de U, la différentielle df_a est un isomorphisme de $\mathbb{E}_n$ vers $\mathbb{F}_n$.

Il est clair que $i) \Rightarrow ii)$. L'implication réciproque est conséquence du théorème local qui suit. Remarquons que, par rapport à l'énoncé de la section précédente, le fait que V soit ouvert et f^{-1} soit continue de V dans U est à présent conséquence de l'inversibilité de df_a en tout point de U.

Ce théorème est remarquable, puisqu'il assure la différentiabilité de f^{-1} sans qu'il soit nécessaire de définir cette application par des formules explicites. L'exercice qui suit en est un bon exemple d'application :

EXERCICE 18-3.26 On munit l'espace vectoriel $\mathbf{E} = \mathcal{S}_n(\mathbf{R})$ des matrices carrées symétriques réelles d'une norme quelconque. Montrer que l'ensemble $U = \mathcal{S}_n^{*+}(\mathbf{R})$ des matrices définies positives est un ouvert de $\mathbf{E}$. Montrer que l'application de U dans lui-même qui, à une matrice M associe son unique racine carrée symétrique positive est un $\mathcal{C}^\infty$-difféomorphisme de U dans lui-même.

18-3.3.3 Complément : théorème d'inversion locale

Un énoncé que l'on démontre en général avant le théorème global est le suivant :

THÉORÈME 18-3.27 Soit $f : U \to \mathbb{F}_n$ une application de classe $\mathcal{C}^k$ définie sur un ouvert U de $\mathbb{E}_n$. On suppose qu'en un point $a \in U$, la différentielle df_a est un isomorphisme. Il existe alors un voisinage ouvert W de a, inclus dans U et un voisinage W' ouvert de $f(a)$ dans $\mathbb{F}_n$ tels que $f(W) = W'$ et

$$f|_W : W \to W'$$

soit un $\mathcal{C}^k$-difféomorphisme.

On dit alors que f définit un $\mathcal{C}^k$-difféomorphisme local en a. Cette fonction définit alors, au voisinage de a, un "bon" changement de variables de classe $\mathcal{C}^k$.

Prenons l'exemple du passage en coordonnées sphériques. On considère l'application f, évidemment de classe $\mathcal{C}^\infty$, définie par

$$f : \mathbb{R}^3 \to \mathbb{R}^3 \quad (r, \theta, \varphi) \mapsto (x, y, z) = (r\cos\theta\cos\varphi, r\sin\theta\cos\varphi, r\sin\varphi)$$

(l'espace d'arrivée est muni de sa structure euclidienne canonique).

La matrice jacobienne de f s'écrit

$$J(f)(r,\theta,\varphi) = \begin{pmatrix} \cos\theta\cos\varphi & -r\sin\theta\cos\varphi & -r\cos\theta\sin\varphi \\ \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\varphi & 0 & r\cos\varphi \end{pmatrix}$$

avec $\det J(f)(r,\theta,\varphi) = r^2\cos\varphi$. Choisissons un triplet $(r_0,\theta_0,\varphi_0) \in \mathbb{R}^3$, avec $r_0 > 0$ et $\varphi_0 \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[$. Le point correspondant $M_0(x_0,y_0,z_0)$ est alors situé hors de l'axe Oz[11]. Comme $r_0^2\cos\varphi_0 \neq 0$, le théorème d'inversion locale s'applique en (r_0,θ_0,φ_0), et on peut trouver un voisinage ouvert W' de M_0 dans $\mathbb{R}^3$, et une application $\mathcal{C}^\infty$

$$\Phi : W' \to \mathbb{R}^3 \quad M(x,y,z) \mapsto (r(x,y,z), \theta(x,y,z), \varphi(x,y,z))$$

telle que pour tout $M(x,y,z) \in W'$, $(r(x,y,z), \theta(x,y,z), \varphi(x,y,z))$ soit un système de coordonnées sphériques de M, avec $\Phi(x_0,y_0,z_0) = (r_0,\theta_0,\varphi_0)$. Ici encore, il n'est nul besoin de connaître explicitement de "formules" définissant Φ pour calculer la matrice jacobienne $J(\Phi)(x_0,y_0,z_0)$. Il suffit d'inverser $J(f)(r_0,\theta_0,\varphi_0)$.

Par contre, au voisinage d'un triplet (r_0,θ_0,φ_0) tel que le point M_0 soit sur l'axe Oz, définir un système de coordonnées polaires $\mathcal{C}^\infty$

$$(r(x,y,z), \theta(x,y,z), \varphi(x,y,z))$$

du point $M(x,y,z)$ n'est pas possible : le projeté orthogonal m_0 de M_0 sur xOy est situé à l'origine, et f n'est pas localement injective : tout point (r_0,α,φ_0) a pour image M_0. Il n'est pas possible de définir θ comme fonction $\mathcal{C}^\infty$ de M au voisinage de M_0.

Démonstration : En fixant des bases au départ et à l'arrivée, on peut supposer $\mathbb{E}_n = \mathbb{F}_n = \mathbb{R}^n$. On peut aussi se ramener au cas où $a = f(a) = 0$, en remplaçant la fonction f par l'application (définie au voisinage de 0) $x \mapsto \varphi(x) = f(a+x) - f(a)$, qui vérifie

$$d\varphi_0 = df_a \in \mathcal{GL}(\mathbb{R}^n)$$

Enfin, en considérant l'application $(d\varphi_0)^{-1} \circ \varphi$, on supposera que f vérifie

$$a = f(a) = 0 \quad \text{et} \quad df_a = id_{\mathbb{R}^n}$$

[11] Réciproquement, tout point $M_0(x_0,y_0,z_0)$ situé hors de Oz admet un système de coordonnées sphériques (r_0,θ_0,φ_0) avec $r_0 > 0$ et $\varphi_0 \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[$: on choisit

$$r_0 = \sqrt{x_0^2 + y_0^2 + z_0^2}$$

puis

$$\varphi_0 = \arcsin \frac{z_0}{\sqrt{x_0^2 + y_0^2 + z_0^2}}$$

et θ_0 tel que $(r_0\cos\varphi_0, \theta_0)$ soit un système de coordonnées polaires du point m_0 projeté orthogonal de M_0 sur le plan xOy. On peut imposer par exemple $\theta_0 \in]-\pi,\pi]$.

La méthode suivie est celle du point fixe : si $y \in \mathbb{R}^n$ est "proche de 0", on cherche à résoudre l'équation $f(x) = y$, avec x "proche de 0". On écrit cette équation sous la forme

$$h_y(x) = x$$

où la fonction h_y est définie sur U par $h_y(x) = y + x - f(x) = y + g(x)$ avec

$$g : U \to \mathbb{R}^n \quad x \mapsto g(x) = x - f(x)$$

On étudie donc les propriétés de la fonction g au voisinage de 0 :

- On a $g(0) = 0$ et $dg_0 = id_{\mathbb{R}^n} - df_0 = 0_{\mathcal{L}(\mathbb{R}^n)}$. Comme g est de classe C^1, on peut trouver un réel $r > 0$ tel que

$$\|x\| \leqslant r \Rightarrow x \in U \text{ et } \|dg_x\| \leqslant \frac{1}{2}$$

(la norme est celle des endomorphismes continus de $\mathbb{R}^n$). La boule fermée $\mathcal{B}(0, r]$ étant convexe, l'inégalité des accroissements finis donne

$$\forall x \in U \quad \|x\| \leqslant r \Rightarrow \|g(x)\| = \|g(x) - g(0)\| \leqslant \frac{1}{2}\|x\| \leqslant \frac{r}{2}$$

et par conséquent

$$g(\mathcal{B}(0, r]) \subset \mathcal{B}\left(0, \frac{r}{2}\right]$$

En manipulant les inégalités strictes, on a une inclusion analogue pour les boules ouvertes.

- A l'aide du théorème du point fixe, nous montrons à présent que tout point $y_0 \in \mathcal{B}\left(0, \frac{r}{2}\right[$ possède un unique antécédent x_0 par f dans la boule ouverte $\mathcal{B}(0, r[$, ce qui revient à résoudre l'équation

$$h_{y_0}(x) = x$$

avec la fonction h_{y_0} introduite plus haut :

Si $\|y_0\| < \frac{r}{2}$, nous avons, pour $\|x\| \leqslant r$,

$$\|h_{y_0}(x)\| = \|y_0 + g(x)\| \leqslant \|y_0\| + \|g(x)\| \leqslant \frac{r}{2} + \frac{r}{2} = r$$

et par conséquent

$$h_{y_0}(\mathcal{B}(0, r]) \subset \mathcal{B}(0, r]$$

Cette boule fermée est une partie complète de $\mathbb{R}^n$, stable par h_{y_0}, application contractante sur cette boule, puisque, par inégalité des accroissements finis

$$x, x' \in \mathcal{B}(0, r] \Rightarrow \|h_{y_0}(x) - h_{y_0}(x')\| = \|g(x) - g(x')\| \leqslant \frac{1}{2}\|x - x'\|$$

La fonction h_{y_0} possède donc un unique point fixe x_0 dans $\mathcal{B}(0, r]$, avec $x_0 = y_0 + g(x_0)$, ce qui entraîne

$$\|x_0\| \leqslant \|y_0\| + \frac{r}{2} < r$$

Nous avons donc bien prouvé :

$$\forall y \in \mathcal{B}\left(0, \frac{r}{2}\right[\quad \exists!\, x \in \mathcal{B}(0, r[\quad f(x) = y$$

- Nous trouvons à présent un voisinage ouvert W de 0 sur lequel f induit une bijection sur $W' = \mathcal{B}\left(0, \frac{r}{2}\right[$. Il suffit de prendre

$$W = f^{-1}\left(\mathcal{B}\left(0, \frac{r}{2}\right[\right) \cap \mathcal{B}(0, r[$$

C'est un voisinage ouvert de 0 (car f est continue), répondant évidemment à la question : $\varphi = f|_W : W \to W' = \mathcal{B}\left(0, \frac{r}{2}\right[$ est une bijection (continue).

- Montrons que φ est bicontinue : soient y et $y' \in W'$, $x = \varphi^{-1}(y)$ et $x' = \varphi^{-1}(y')$. On a $x, x' \in \mathcal{B}(0, r[$, avec

$$x - x' = g(x) - g(x') + f(x) - f(x') = g(x) - g(x') + y - y'$$

Comme g est $\frac{1}{2}$-lipschitzienne sur $\mathcal{B}(0, r[$, on obtient, par inégalité triangulaire

$$\|x - x'\| \leqslant \frac{1}{2}\|x - x'\| + \|y - y'\|$$

ce qui donne finalement

$$\forall y, y' \in W' \quad \left\|\varphi^{-1}(y) - \varphi^{-1}(y')\right\| \leqslant 2\,\|y - y'\|$$

La continuité de φ^{-1} est ainsi prouvée.

- Enfin, si on a pris la peine de choisir r suffisamment petit, on est assuré (par continuité) que

$$\forall x \in W \quad \det(d\varphi_x) \neq 0$$

Le corollaire **18-2**.15 nous montre alors que φ^{-1} est de classe C^1. On montre aisément (par récurrence sur k) qu'elle est C^k si f l'est aussi. ■

EXERCICE 18-3.28 Avec le théorème d'inversion locale, démontrer le théorème d'inversion globale. Montrer également qu'une application C^1 d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}^n$ dont la différentielle est en tout point inversible transforme tout ouvert de U en un ouvert de $\mathbb{R}^n$.

18-3.4 Théorème des fonctions implicites

18-3.4.1 Rappel : résolution d'un système linéaire

L'idée générale du calcul différentiel est l'approximation locale d'une fonction non linéaire $h \mapsto f(a+h) - f(a)$ par une fonction linéaire $h \mapsto df_a(h)$. Commençons donc par rappeler le résultat concernant la résolution d'un système linéaire d'équations (voir le théorème **4-5**.1) :

Nous représentons ici les vecteurs de $\mathbb{R}^p$, $\mathbb{R}^q$ et $\mathbb{R}^{p+q}$ indifféremment sous forme de vecteurs lignes ou colonnes. On considère une matrice $M \in \mathcal{M}_{q,p+q}(\mathbb{R})$, que l'on écrit par blocs

$$M = (A \mid B)$$

avec $A \in \mathcal{M}_{q,p}(\mathbb{R})$ et $B \in \mathcal{M}_q(\mathbb{R})$, et un vecteur $C \in \mathbb{R}^q$. On définit avec ces données une application f de $\mathbb{R}^{p+q}$ vers $\mathbb{R}^q$ par

$$(X,Y) \mapsto f(X,Y) = M\begin{pmatrix} X \\ Y \end{pmatrix} - C = AX + BY - C$$

les vecteurs X et Y étant respectivement des vecteurs colonnes de $\mathbb{R}^p$ et $\mathbb{R}^q$, et l'espace $\mathbb{R}^{p+q}$ étant évidemment identifié à $\mathbb{R}^p \times \mathbb{R}^q$. On cherche à résoudre l'équation $f(X,Y) = 0$, ce qui s'écrit, avec des notations évidentes

$$\begin{cases} m_{1,1}x_1 + \cdots + m_{1,p}x_p + m_{1,p+1}y_1 + \cdots + m_{1,p+q}y_q = c_1 \\ \vdots \\ m_{i,1}x_1 + \cdots + m_{i,p}x_p + m_{i,p+1}y_1 + \cdots + m_{i,p+q}y_q = c_i \\ \vdots \\ m_{q,1}x_1 + \cdots + m_{q,p}x_p + m_{q,p+1}y_1 + \cdots + m_{q,p+q}y_q = c_q \end{cases}$$

Il s'agit de résoudre un système linéaire à $p+q$ inconnues

$$(x_1, \ldots, x_p, x_{p+1} = y_1, \ldots, x_{p+q} = y_q)$$

comportant q équations. On suppose que ce système est de rang q, et plus précisément que les inconnues $(y_1, \ldots, y_q)$ peuvent être utilisées comme inconnues principales (*cf.* section **4-5.2**), ce qui veut dire que

$$\det B = \begin{vmatrix} m_{1,p+1} & \cdots & m_{1,p+q} \\ \vdots & & \vdots \\ m_{q,p+1} & \cdots & m_{q,p+q} \end{vmatrix} \neq 0 \quad \text{c'est-à-dire } B \text{ est inversible.}$$

On sait alors que, pour toutes valeurs des inconnues secondaires $X \in \mathbb{R}^p$, le système possède une solution (X,Y) donnée par

$$f(X,Y) = 0 \Leftrightarrow AX + BY = C \Leftrightarrow Y = B^{-1}C - B^{-1}AX$$

Y étant obtenu en résolvant un système de Cramer.

Remarquons que, puisque f est affine de $\mathbb{R}^{p+q}$ dans $\mathbb{R}^q$, la matrice M est matrice jacobienne (constante) de f, et le bloc B pourra être interprété, conformément à la définition de la section suivante, comme matrice d'une différentielle

partielle (par rapport à la variable Y). On a donc, dans le cas d'un système linéaire du type précédent, Y comme fonction "explicite" de X. Remarquons également que, dans ce cas, la correspondance $X \mapsto Y$ est affine, et que $N = -B^{-1}A$ en est la matrice jacobienne.

Dans le cas d'une équation non linéaire, le théorème qui suit donnera une condition suffisante pour que, localement, la relation $f(X,Y) = 0$ définisse Y comme fonction "implicite" de X.

18-3.4.2 Théorème des fonctions implicites

DÉFINITION 18-3.29 *Soit f une application de classe C^1 d'un ouvert U de $\mathbb{R}^{p+q} \approx \mathbb{R}^p \times \mathbb{R}^q$ à valeurs dans $\mathbb{R}^m$ (ou à valeurs dans un espace normé de dimension m ; on se ramène à $\mathbb{R}^m$par le choix d'une base) :*

$$f : (x,y) \mapsto f(x,y) = (f_1(x,y), \dots, f_m(x,y))$$

Soit $(a,b) \in U$. On appelle différentielle partielle de f en (a,b) par rapport à la deuxième variable la différentielle en b de l'application (définie au voisinage de b)

$$y \mapsto f(a,y)$$

On la note $d_2 f_{(a,b)} \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^m)$. On définirait de même la différentielle par rapport à la première variable $d_1 f(a,b) \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^m)$ en différentiant en a l'application $x \mapsto f(x,b)$.

Comme f est différentiable en (a,b), on a, en munissant par exemple les espaces $\mathbb{R}^p$, $\mathbb{R}^q$ et $\mathbb{R}^{p+q}$ des normes $\| \|_\infty$:

$$f(a+h, b+k) = f(a,b) + df_{(a,b)}(h,k) + \max(\|h\|, \|k\|)\,\varepsilon(h,k)$$

avec $\lim\limits_{(0,0)} \varepsilon(h,k) = 0$, ce qui donne évidemment

$$\begin{cases} f(a, b+k) = f(a,b) + df_{(a,b)}(0,k) + \|k\|\,\varepsilon(0,k) \\ f(a+h, b) = f(a,b) + df_{(a,b)}(h,0) + \|h\|\,\varepsilon(h,0) \end{cases}$$

ce qui prouve l'existence des différentielles partielles, avec simplement

$$d_2 f_{(a,b)}(k) = df_{(a,b)}(0,k) \quad \text{et} \quad d_1 f_{(a,b)}(h) = df_{(a,b)}(h,0)$$

ce qui donne aussi, par linéarité de $df_{(a,b)}$

$$df_{(a,b)}(h,k) = d_1 f_{(a,b)}(h) + d_2 f_{(a,b)}(k)$$

Matriciellement, lorsqu'on travaille dans les bases canoniques, on obtient une décomposition par blocs de la jacobienne de f en (a,b) :

$$J(f)(a,b) = (A \mid B)$$

avec

$$B = \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1}(a,b) & \cdots & \dfrac{\partial f_1}{\partial y_q}(a,b) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial y_1}(a,b) & \cdots & \dfrac{\partial f_m}{\partial y_q}(a,b) \end{pmatrix} \in \mathcal{M}_{m,q}(\mathbb{R})$$

qui est la matrice canoniquement associée à $d_2f_{(a,b)} \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^m)$.

Ces notations étant acquises, énonçons le théorème des fonctions implicites :

THÉORÈME 18-3.30 Soit U un ouvert de $\mathbb{R}^p \times \mathbb{R}^q$ et f une application de classe C^k de U dans $\mathbb{R}^q$:

$$f : U \ni (x,y) \mapsto f(x,y) = (f_1(x,y), \ldots, f_q(x,y))$$

On suppose qu'il existe un point $M_0 = (a,b) \in U$ tel que

- $f(a,b) = 0$
- **La différentielle partielle $d_2f(a,b)$ est un automorphisme de $\mathbb{R}^q$, ce qui se traduit par la non-nullité du jacobien partiel :**

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial y_1}(a,b) & \cdots & \dfrac{\partial f_1}{\partial y_q}(a,b) \\ \vdots & & \vdots \\ \dfrac{\partial f_q}{\partial y_1}(a,b) & \cdots & \dfrac{\partial f_q}{\partial y_q}(a,b) \end{vmatrix} \neq 0$$

Il existe alors un voisinage ouvert V de a dans $\mathbb{R}^p$, un voisinage ouvert W de b dans $\mathbb{R}^q$ et une application $\varphi : V \to W$ de classe C^k tels que

- $V \times W \subset U$
- $\forall (x,y) \in V \times W \quad f(x,y) = 0 \Leftrightarrow y = \varphi(x)$

On a évidemment $\varphi(a) = b$, et la différentielle de φ en a vaut

$$d\varphi_a = -\left[d_2f_{(a,b)}\right]^{-1} \circ d_1f_{(a,b)}$$

Ce théorème est local : on n'a pas d'information *a priori* sur la taille des ouverts V et W. Ils peuvent être très "petits".
Si on considère

$$S = \{(x,y) \in \mathbb{R}^{p+q} \mid f(x,y) = 0\}$$

et si les hypothèses du théorème des fonctions implicites sont vérifiées en un point $M_0 = (a,b)$ de S, l'ensemble S est, au voisinage de (a,b) (cela signifiant qu'on ne considère que la trace de S sur $V \times W$), le graphe d'une application $\varphi : V \to W$ de classe C^k :

$$\begin{aligned} &\{(x,y) \in \mathbb{R}^{p+q} \mid x \in V,\ y \in W \text{ et } f(x,y) = 0\} \\ &\qquad = S \cap (V \times W) = \{(x, \varphi(x)),\ x \in V\} \end{aligned}$$

On dit qu'au voisinage de (a,b), la relation $f(x,y) = 0$ définit y comme fonction implicite[12] de x. En particulier

$$\forall x \in V \quad \exists!\, y \in W \quad f(x,y) = 0$$

[12] Si l'on préfère travailler avec des équations scalaires, le système de q équations scalaires

$$\begin{cases} f_1(x_1, \ldots, x_p, y_1, \ldots, y_q) = 0 \\ \qquad \vdots \\ f_q(x_1, \ldots, x_p, y_1, \ldots, y_q) = 0 \end{cases}$$

Démonstration : Considérons l'application de U dans $\mathbb{R}^{p+q}$ définie par

$$\Phi : (x,y) \mapsto (x, f(x,y))$$

Cette application est évidemment de classe C^k, puisque ses composantes dans la base canonique de $\mathbb{R}^{p+q}$ le sont. Sa matrice jacobienne en (a,b) vaut

$$J(\Phi)(a,b) = \begin{pmatrix} I_p & 0 \\ \begin{array}{|ccc|} \hline \frac{\partial f_1}{\partial x_1}(a,b) & \cdots & \frac{\partial f_1}{\partial x_p}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial x_1}(a,b) & \cdots & \frac{\partial f_q}{\partial x_p}(a,b) \\ \hline \end{array} & \begin{array}{|ccc|} \hline \frac{\partial f_1}{\partial y_1}(a,b) & \cdots & \frac{\partial f_1}{\partial y_q}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial y_1}(a,b) & \cdots & \frac{\partial f_q}{\partial y_q}(a,b) \\ \hline \end{array} \end{pmatrix}$$

et est inversible. D'après le théorème d'inversion locale (section **18-3.3.3**), Φ est un C^k-difféomorphisme local en (a,b). Plus précisément, il existe un voisinage ouvert V_1 de (a,b) dans U et un voisinage ouvert W_1 de $\Phi(a,b) = (a,0)$ dans $\mathbb{R}^{p+q}$ tel que $\Phi|_{V_1} : V_1 \to W_1$ soit un C^k-difféomorphisme. Comme V_1 est un voisinage de (a,b), il contient un sous-ensemble de la forme $V \times W$, avec $V \in \mathcal{V}(a)$ et $W \in \mathcal{V}(b)$ ouverts. Comme $\Phi|_{V_1}$ est un homéomorphisme $V_2 = \Phi(V \times W)$ est un ouvert de W_1 donc de $\mathbb{R}^{p+q}$ et

$$\Phi|_{V \times W} : V \times W \to V_2$$

est un C^k difféomorphisme. Si on note $\Psi = (\Phi|_{V \times W})^{-1}$ le difféomorphisme réciproque, comme

$$\forall (x,y) \in V \times W \quad \Phi(x,y) = (x, f(x,y))$$

on a

$$\forall z \in \mathbb{R}^p \quad \forall t \in \mathbb{R}^q \quad (z,t) \in V_2 \Rightarrow z \in V \text{ et } \Psi(z,t) = (z, g(z,t))$$

avec g de classe C^k de V_2 dans W. On a, pour $x \in V$ et $y \in W$

$$f(x,y) = 0 \Leftrightarrow \Phi(x,y) = (x,0) \Leftrightarrow (x,y) = \Psi(x,0) = (x, g(x,0))$$

---

est équivalent, pour $(x_1, \ldots, x_p) \in V$ et $(y_1, \ldots, y_q) \in W$ à

$$\begin{cases} y_1 = \varphi_1(x_1, \ldots, x_p) \\ \vdots \\ y_q = \varphi_q(x_1, \ldots, x_p) \end{cases}$$

Le système peut donc être localement résolu (pas explicitement en général) en $(y_1, \ldots, y_q)$: ces q inconnues sont fonctions implicites des p autres. Les applications φ_i, composantes de φ, sont aussi de classe C^k.

ce qui démontre l'existence de la fonction φ de l'énoncé du théorème, avec

$$\varphi : V \to W \text{ définie par } \varphi(x) = g(x, 0)$$

Comme

$$\forall x \in V \quad f(x, \varphi(x)) = 0$$

la différentielle de l'application $V \ni x \mapsto f(x, \varphi(x))$ est partout nulle. Si ψ est l'application : $V \to V \times W$ définie par $\psi(x) = (x, \varphi(x))$, on a évidemment

$$\forall h \in \mathbb{R}^p \quad d\psi_x(h) = (h, d\varphi_x(h))$$

Le théorème de différentiation d'une fonction composée donne alors

$$\forall x \in V \quad \forall h \in \mathbb{R}^p \quad df_{(x,\varphi(x))}(h, d\varphi_x(h)) = 0_{\mathbb{R}^q}$$

ce qui peut s'écrire aussi, avec les différentielles partielles

$$\forall x \in V \quad \forall h \in \mathbb{R}^p \quad d_1 f_{(x,\varphi(x))}(h) + d_2 f_{(x,\varphi(x))}(d\varphi_x(h)) = 0_{\mathbb{R}^q}$$

soit enfin

$$\forall x \in V \quad d_1 f_{(x,\varphi(x))} + d_2 f_{(x,\varphi(x))} \circ d\varphi_x = 0 \qquad (**)$$

En particulier pour $x = a$, on obtient $d_1 f_{(a,b)} + d_2 f_{(a,b)} \circ d\varphi_a = 0$, ce qui donne, puisque $d_2 f_{(a,b)}$ est un isomorphisme

$$d\varphi_a = -\left[d_2 f_{(a,b)}\right]^{-1} \circ d_1 f_{(a,b)} \quad \blacksquare \qquad (***)$$

REMARQUE 18-3.31 On a supposé que $d_2 f_{(a,b)}$ est un automorphisme de $\mathbb{R}^q$, ce qui se traduit par

$$\det\left(d_2 f_{(a,b)}\right) \neq 0$$

Par continuité de l'application $V \ni x \mapsto \det\left(d_2 f_{(x,\varphi(x))}\right)$, il existe un voisinage V' de a inclus dans V sur lequel $\det\left(d_2 f_{(x,\varphi(x))}\right) \neq 0$. On a alors, d'après l'égalité (**)

$$\forall x \in V' \quad d\varphi_x = -\left[d_2 f_{(x,\varphi(x))}\right]^{-1} \circ d_1 f_{(x,\varphi(x))}$$

Nous verrons sur des exemples simples, en petites dimensions, que cette formule se retrouve par des calculs simples de dérivées partielles, le théorème des fonctions implicites assurant le caractère $\mathcal{C}^k$ de φ et permettant ainsi le calcul de dérivées partielles de fonctions composées.

REMARQUE 18-3.32 Le fait de travailler avec une équation dont le second membre est 0 n'est pas restrictif : si (x_0, y_0) est un point de U où la différentielle partielle par rapport à y est un automorphisme de $\mathbb{R}^q$, pour résoudre, au voisinage de ce point l'équation

$$f(x, y) = f(x_0, y_0)$$

on appliquera tout simplement le théorème des fonctions implicites à la fonction g définie sur U par

$$g(x, y) = f(x, y) - f(x_0, y_0)$$

qui a évidemment la même différentielle que f.

18-3.4.3 Exemples en petites dimensions

"Courbe" d'équation $f(x,y)=0$

Dans le plan rapporté à un repère $(O,\vec{\imath},\vec{\jmath})$, on considère, f étant une application de classe C^k d'un ouvert de $\mathbb{R}^2$ dans $\mathbb{R}$, l'ensemble

$$\Gamma=\{M(x,y)\mid f(x,y)=0\}$$

DÉFINITION 18-3.33 *Un point $M_0(x_0,y_0)\in\Gamma$ est dit* ***régulier*** *si et seulement si*

$$df_{M_0}\neq 0$$

c'est-à-dire si l'une au moins des dérivées partielles de f en M_0 est non nulle

$$\left(\frac{\partial f}{\partial x}(x_0,y_0),\frac{\partial f}{\partial y}(x_0,y_0)\right)\neq(0,0)$$

Supposons par exemple que $\dfrac{\partial f}{\partial y}(x_0,y_0)\neq 0$. Le théorème des fonctions implicites montre alors qu'il existe un voisinage ouvert $\mathcal{V}\times\mathcal{W}$ de M_0 tel que $\Gamma\cap(\mathcal{V}\times\mathcal{W})$ soit le support d'un arc paramétré de classe C^k

$$\Gamma\cap(\mathcal{V}\times\mathcal{W})=\{M(x,\varphi(x)),\ x\in\mathcal{V}\}$$

avec φ fonction de classe C^k de $\mathcal{V}$ dans $\mathcal{W}$ (en diminuant si besoin l'ensemble $\mathcal{V}$, on peut supposer que $\mathcal{V}$ est un intervalle ouvert, par exemple de la forme $]x_0-\varepsilon,x_0+\varepsilon[$). On obtient en fait la représentation graphique de la fonction φ, dans le repère $(O,\vec{\imath},\vec{\jmath})$. Cet arc présente en M_0 une tangente $\mathcal{T}_0$, d'équation

$$Y-y_0=\varphi'(x_0)(X-x_0)$$

La dérivée de la fonction implicite s'obtient par la formule (***), qui se retrouve en dérivant la fonction constante $x\mapsto f(x,\varphi(x))$ soit

$$\forall x\in\mathcal{V}\quad \frac{\partial f}{\partial x}(x,\varphi(x))+\frac{\partial f}{\partial y}(x,\varphi(x))\,\varphi'(x)=0$$

donc

$$\varphi'(x_0)=-\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}(x_0,y_0)$$

En chassant le dénominateur, on obtient l'équation de la tangente $\mathcal{T}_0$ en M_0

$$\frac{\partial f}{\partial x}(M_0).(X-x_0)+\frac{\partial f}{\partial y}(M_0).(Y-y_0)=0$$

On s'aperçoit que les deux variables x et y jouent des rôles tout à fait symétriques dans cette équation, et qu'on serait arrivé au même résultat si l'on avait supposé

$\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$ (Γ serait alors, au voisinage de M_0, paramétré par y). On a dans les deux cas

$$M(X,Y) \in \mathcal{T}_0 \Leftrightarrow \overrightarrow{M_0M} \in \ker df_{M_0}$$

La direction de la tangente est donc le noyau de la différentielle de f en M_0. Si le plan est euclidien, cette différentielle s'écrit

$$df_{M_0} = \left\langle \overrightarrow{\text{grad} f}(M_0), \bullet \right\rangle$$

et le vecteur $\overrightarrow{\text{grad} f}(M_0)$ est normal à Γ en M_0 :

$$\mathcal{T}_0 = M_0 + \left[\overrightarrow{\text{grad} f}(M_0)\right]^{\perp}$$

EXERCICE 18-3.34 Montrer qu'il existe une unique fonction $\varphi : \mathbb{R}^+ \to \mathbb{R}$ vérifiant

$$\forall x \geqslant 0 \quad [\varphi(x)]^3 + x\varphi(x) = e^x$$

Montrer que φ est $\mathcal{C}^\infty$. Etudier ses variations. En donner un développement limité d'ordre 3 en 0. Donner un équivalent simple de $\varphi(x)$ pour $x \to +\infty$.

"Surface" d'équation $f(x,y,z) = 0$

Dans un espace affine de dimension 3 rapporté à un repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$, on considère à présent

$$S = \{M(x,y,z) \mid f(x,y,z) = 0\}$$

f étant une fonction numérique de classe $\mathcal{C}^k$ définie sur un ouvert de $\mathbb{R}^3$. Comme précédemment

DÉFINITION 18-3.35 *Un point $M_0(x_0, y_0, z_0) \in S$ est dit* ***régulier*** *si et seulement si*

$$df_{M_0} \neq 0$$

c'est-à-dire si l'une au moins des dérivées partielles de f en M_0 est non nulle

$$\left(\frac{\partial f}{\partial x}(x_0, y_0, z_0), \frac{\partial f}{\partial y}(x_0, y_0, z_0), \frac{\partial f}{\partial z}(x_0, y_0, z_0)\right) \neq (0,0,0)$$

Supposons par exemple que $\frac{\partial f}{\partial z}(M_0) \neq 0$. Le théorème des fonctions implicites montre alors qu'il existe un voisinage ouvert $\mathcal{V}$ de (x_0, y_0) et un voisinage ouvert $\mathcal{W}$ de z_0 tel que

$$S \cap (\mathcal{V} \times \mathcal{W}) = \{M(x, y, \varphi(x,y)), \quad (x,y) \in \mathcal{V}\}$$

avec φ fonction de classe $\mathcal{C}^k$ de $\mathcal{V}$ dans $\mathcal{W}$. $S \cap (\mathcal{V} \times \mathcal{W})$ est en fait un "morceau de surface" d'équation[13] $z = \varphi(x,y)$ dans le repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$. Conformément à

[13]On dira aussi le support d'une nappe paramétrée de classe $\mathcal{C}^k$

$$\mathcal{V} \ni (x,y) \mapsto M(x, y, \varphi(x,y))$$

où (x,y) est le paramètre.

l'étude de la section **18-2.3.2**, cette surface possède en M_0 un plan tangent $\mathcal{T}_0$, d'équation

$$(Z - z_0) = \frac{\partial \varphi}{\partial x}(x_0, y_0)(X - x_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0)(Y - y_0)$$

Ici encore les dérivées partielles de la fonction implicite φ peuvent s'obtenir en dérivant partiellement la fonction constante $(x, y) \mapsto f(x, y, \varphi(x, y))$ soit

$$\forall (x, y) \in \mathcal{V} \quad \begin{cases} \dfrac{\partial f}{\partial x}(x, y, \varphi(x, y)) + \dfrac{\partial f}{\partial z}(x, y, \varphi(x, y)) \dfrac{\partial \varphi}{\partial x}(x, y) = 0 \\ \dfrac{\partial f}{\partial y}(x, y, \varphi(x, y)) + \dfrac{\partial f}{\partial z}(x, y, \varphi(x, y)) \dfrac{\partial \varphi}{\partial z}(x, y) = 0 \end{cases}$$

soit

$$\frac{\partial \varphi}{\partial x}(x_0, y_0) = -\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial z}}(M_0) \quad \text{et} \quad \frac{\partial \varphi}{\partial y}(x_0, y_0) = -\frac{\dfrac{\partial f}{\partial y}}{\dfrac{\partial f}{\partial z}}(M_0)$$

En chassant le dénominateur, on obtient l'équation du plan tangent $\mathcal{T}_0$ en M_0

$$\frac{\partial f}{\partial x}(M_0) . (X - x_0) + \frac{\partial f}{\partial y}(M_0) . (Y - y_0) + \frac{\partial f}{\partial z}(M_0) . (Z - z_0) = 0$$

Les trois variables x, y et z jouent des rôles tout à fait symétriques dans cette équation.

$$M(X, Y, Z) \in \mathcal{T}_0 \Leftrightarrow \overrightarrow{M_0 M} \in \ker df_{M_0}$$

Cette condition était prévisible : si $I \ni t \mapsto M(t)$ est un arc paramétré de classe $\mathcal{C}^1$ régulier tracé sur S, on a par définition de s

$$\forall t \in I \quad f(M(t)) = 0$$

ce qui donne par dérivation

$$\forall t \in I \quad df_{M(t)}\left(\vec{M}'(t)\right) = 0$$

et si, pour $t = 0$ on a $M(0) = M_0$, on a bien $\vec{M}'(0) \in \ker df_{M_0}$.

Si l'espace est euclidien, le vecteur $\overrightarrow{\text{grad} f}(M_0)$ est normal à S en M_0 :

$$\mathcal{T}_0 = M_0 + \left[\overrightarrow{\text{grad} f}(M_0)\right]^{\perp}$$

EXERCICE 18-3.36 Soit S l'ensemble des points de $\mathbf{R}^3$ vérifiant l'équation

$$x^3 + y^3 + z^3 - xyz - 2 = 0$$

Prouver l'existence d'un plan tangent à S en $M_0(1, 1, 1)$. Donner l'équation de ce plan et déterminer la position de S par rapport à ce plan tangent au voisinage de M_0.

Intersection de deux surfaces "transverses" d'équations

$$f(x,y,z)=0 \text{ et } g(x,y,z)=0$$

Dans un espace affine de dimension 3 rapporté à un repère $\left(O,\vec{\imath},\vec{\jmath},\vec{k}\right)$, on considère à présent

$$S_f=\{M(x,y,z)\mid f(x,y,z)=0\} \text{ et } S_g=\{M(x,y,z)\mid g(x,y,z)=0\}$$

f et g étant deux fonctions numériques de classe C^k définies sur un ouvert de $\mathbb{R}^3$ et l'on étudie l'intersection $S_f\cap S_g$. On suppose que le point $M_0(x_0,y_0,z_0)\in S_f\cap S_g$ est à la fois **régulier pour** S_f **et** S_g. On suppose aussi que les deux plans tangents à S_f et S_g en M_0 sont distincts. Comme ces deux plans passent par M_0, cela revient à dire que les directions de ces plans sont distinctes, soit

$$\ker df_{M_0}\neq\ker dg_{M_0}\Leftrightarrow \operatorname{rang}\{df_{M_0},dg_{M_0}\}=2$$

Ceci revient à dire que

$$\operatorname{rang}\begin{pmatrix}\dfrac{\partial f}{\partial x}(M_0) & \dfrac{\partial f}{\partial y}(M_0) & \dfrac{\partial f}{\partial z}(M_0)\\ \dfrac{\partial g}{\partial x}(M_0) & \dfrac{\partial g}{\partial y}(M_0) & \dfrac{\partial g}{\partial z}(M_0)\end{pmatrix}=2$$

donc qu'un déterminant d'ordre 2 extrait de cette matrice est non nul.

Supposons par exemple

$$\begin{vmatrix}\dfrac{\partial f}{\partial y}(M_0) & \dfrac{\partial f}{\partial z}(M_0)\\ \dfrac{\partial g}{\partial y}(M_0) & \dfrac{\partial g}{\partial z}(M_0)\end{vmatrix}\neq 0$$

Le théorème des fonctions implicites montre alors qu'il existe un voisinage ouvert $\mathcal{V}$ de x_0 et un voisinage ouvert $\mathcal{W}$ de (y_0,z_0) tel que

$$(S_f\cap S_g)\cap(\mathcal{V}\times\mathcal{W})=\{M(x,\varphi_1(x),\varphi_2(x)),\ x\in\mathcal{V}\}$$

avec $\varphi=(\varphi_1,\varphi_2)$ fonction de classe C^k de $\mathcal{V}$ dans $\mathcal{W}$. L'intersection

$$(S_f\cap S_g)\cap(\mathcal{V}\times\mathcal{W})$$

est donc le support d'un arc paramétré γ de classe C^k

$$x\mapsto M(x,\varphi_1(x),\varphi_2(x))$$

dans le repère $\left(O,\vec{\imath},\vec{\jmath},\vec{k}\right)$. Comme précédemment, on pourrait placer la tangente à cet arc en M_0 en déterminant le vecteur dérivé

$$\overrightarrow{M'}(x_0)=\vec{\imath}+\varphi_1'(x_0)\vec{\jmath}+\varphi_2'(x_0)\vec{k}$$

les dérivées des fonctions φ_1 et φ_2 pouvant s'obtenir en résolvant le système d'équations obtenues en dérivant les fonctions constantes $x\mapsto f(x,\varphi_1(x),\varphi_2(x))$ et $x\mapsto f(x,\varphi_1(x),\varphi_2(x))$. On peut aussi remarquer que, γ étant un arc régulier tracé sur S_f et S_g, sa tangente en M_0 est dans le plan tangent à S_f en M_0 et qu'il en est de même en travaillant avec S_g. **La tangente à l'arc $\gamma=S_f\cap S_g$ en M_0 est donc l'intersection des plans tangents à S_f et S_g en M_0.**

18-3.4.4 Extrema liés

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une application de classe $\mathcal{C}^1$. On sait qu'un extremum local éventuel de f est à chercher parmi ses points critiques. On cherche à résoudre ici un autre type de problème : on veut maximiser (ou minimiser) la fonction $(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ sous la contrainte $g(x_1, \ldots, x_n) = 0$. Plus précisément, on se donne une autre fonction numérique g de classe $\mathcal{C}^1$ sur U, on considère "l'hypersurface"

$$S_g = \{(x_1, \ldots, x_n) \in U \mid g(x_1, \ldots, x_n) = 0\}$$

(que l'on suppose non vide!) et on cherche les extrema locaux de la restriction $f|_{S_g}$ de f à S_g. Comme S_g n'est pas un ouvert de $\mathbb{R}^n$, ces extrema locaux ne sont pas nécessairement des points critiques de f. Nous allons voir que, moyennant une hypothèse supplémentaire sur S_g, le calcul différentiel permet d'obtenir une **condition nécessaire** pour qu'un point $M_0 \in S_g$ soit extremum local de f sur S_g.

On suppose que tout point $M_0 \in S_g$ est régulier, c'est-à-dire

$$\forall\, M_0 \in S_g \quad dg_{M_0} \neq 0$$

Cela signifie qu'une au moins des dérivées partielles de g est non nulle en M_0, et, comme nous l'avons fait pour les courbes de $\mathbb{R}^2$ ou les surfaces de $\mathbb{R}^3$, nous supposons sans nuire à la généralité que

$$\frac{\partial g}{\partial x_n}(M_0) \neq 0$$

Le théorème des fonctions implicites permet alors de décrire S_g au voisinage de $M_0(x_1^0, \ldots, x_n^0)$ comme étant le graphe d'une fonction φ de classe $\mathcal{C}^1$ définie au voisinage du point m_0 de $\mathbb{R}^{n-1}$ de coordonnées $\left(x_1^0, \ldots, x_{n-1}^0\right)$

$$(x_1, \ldots, x_{n-1}) \mapsto x_n = \varphi(x_1, \ldots, x_{n-1})$$

Si $f|_{S_g}$ est extrémale en M_0, la fonction h de classe $\mathcal{C}^1$ définie sur un voisinage ouvert de m_0

$$h : (x_1, \ldots, x_{n-1}) \mapsto f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1}))$$

est extrémale en m_0, ce qui entraîne que m_0 est point critique de h :

$$\forall\, i = 1 \ldots n-1 \quad 0 = \frac{\partial h}{\partial x_i}(m_0) = \frac{\partial f}{\partial x_i}(M_0) + \frac{\partial f}{\partial x_n}(M_0) \times \frac{\partial \varphi}{\partial x_i}(m_0)$$

Le théorème des fonctions implicites donne

$$\frac{\partial \varphi}{\partial x_i}(m_0) = -\frac{\dfrac{\partial g}{\partial x_i}}{\dfrac{\partial g}{\partial x_n}}(M_0)$$

et on obtient donc

$$\forall i = 1 \ldots n-1 \quad \begin{vmatrix} \dfrac{\partial f}{\partial x_i} & \dfrac{\partial g}{\partial x_i} \\ \dfrac{\partial f}{\partial x_n} & \dfrac{\partial g}{\partial x_n} \end{vmatrix} (M_0) = 0$$

ces égalités montrant que les vecteurs lignes

$$\left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right)(M_0) \text{ et } \left(\frac{\partial g}{\partial x_1}, \cdots, \frac{\partial g}{\partial x_n}\right)(M_0)$$

sont colinéaires. On en déduit la

PROPOSITION 18-3.37 Avec les hypothèses qui précèdent, si $M_0 \in S_g$ est extremum local de $f|_{S_g}$, alors

$$\exists\, \lambda \in \mathbb{R} \quad df_{M_0} = \lambda dg_{M_0}$$

Un point M_0 vérifiant cette condition (nécessaire) est fort logiquement appelé point critique de $f|_{S_g}$ (le cas $\lambda = 0$ correspondrait à un point critique de f situé sur S_g).

EXEMPLE 18-3.38 Déterminer les extrema de la fonction $f(x,y,z) = x+y+z$ sur la sphère unité S de $\mathbb{R}^3$ euclidien canonique. On prend ici

$$g(x,y,z) = x^2 + y^2 + z^2 - 1$$

Comme $\overrightarrow{\text{grad}}\, g(x,y,z) = 2(x,y,z)$ ne s'annule en aucun point de S, on peut appliquer ce qui précède : si $M_0(x_0, y_0, z_0)$ est extremum local de $f|_S$

$$\overrightarrow{\text{grad}}\, f(x,y,z) = (1,1,1) \text{ est colinéaire à } (x,y,z)$$

ce qui donne $x = y = z = \pm\dfrac{\sqrt{3}}{3}$ (on cherche des points sur la sphère !). Il y a donc deux points critiques pour $f|_S$. Comme souvent, c'est un argument de compacité qui assure que ces points critiques sont effectivement des extrema : la fonction numérique continue f atteint son maximum et son minimum sur le compact S

$$\max_S f = f\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right) = \sqrt{3} \text{ alors que } \min_S f = -\sqrt{3}$$

Remarquons que le recours au calcul différentiel n'était pas indispensable ici : par inégalité de Schwarz

$$|f(x,y,z)| \leqslant \sqrt{3}\sqrt{x^2+y^2+z^2}$$

avec égalité si et seulement si les vecteurs $(1,1,1)$ et (x,y,z) sont colinéaires.

EXEMPLE 18-3.39 Nous avons, au chapitre **14**, montré qu'un endomorphisme auto-adjoint u d'un espace vectoriel euclidien $\mathbb{E}_n$ était diagonalisable dans une base orthonormale, le point crucial de la démonstration étant l'existence d'une valeur propre (réelle!) pour u. Nous avons donné une démonstration matricielle de ce résultat, en utilisant en fait la structure hermitienne canonique de $\mathbb{C}^n$.

Voici une démonstration basée sur un résultat d'optimisation : sur la sphère unité de $\mathbb{E}_n$, montrer que $f : x \mapsto \langle u(x), x\rangle$ atteint son maximum en un point x_0. Montrer que

$$\forall h \in \mathbb{E}_n \quad df_{x_0}(h) = 2\langle u(x_0), h\rangle$$

et en déduire qu'un point de la sphère où f atteint son maximum est un vecteur propre de u.

EXERCICE 18-3.40 Dans le cas $n = 2$, en travaillant avec les "lignes de niveau" de f définies pour $k \in \mathbb{R}$ par

$$\gamma_k = \{M(x, y) \in U \mid f(x, y) = k\}$$

et avec la courbe

$$\gamma'_0 = S_g = \{M(x, y) \in U \mid g(x, y) = 0\}$$

donner une interprétation géométrique de la condition

$$\exists \lambda \in \mathbb{R} \quad df_{M_0} = \lambda dg_{M_0}$$

nécessaire pour que M_0 soit point critique de $f|_{S_g}$. Expliquer.

EXERCICE 18-3.41 Cas de p contraintes "indépendantes".

On se donne toujours $f : U \to \mathbb{R}$ $(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$ et p applications de classe C^1 $(g^i)_{1\leqslant i\leqslant p}$ de U dans $\mathbb{R}$, avec $p < n$. On cherche cette fois à optimiser f sous les contraintes

$$\begin{cases} g^1(x_1, \dots, x_n) = 0 \\ \vdots \\ g^p(x_1, \dots, x_n) = 0 \end{cases}$$

On cherche donc les extrema locaux de $f|_{S_{g^1}\cap S_{g^2}\cdots\cap S_{g^p}}$. On suppose bien entendu $S_{g^1} \cap S_{g^2} \cdots \cap S_{g^p} \neq \emptyset$, et on suppose de plus que tout point $M_0 \in S_{g^1} \cap S_{g^2} \cdots \cap S_{g^p}$ est régulier, c'est-à-dire que

$$\text{rang}\left(dg^1_{M_0}, \dots, dg^p_{M_0}\right) = p$$

En utilisant le théorème des fonctions implicites, montrer qu'au voisinage de M_0, l'ensemble $S_{g^1} \cap S_{g^2} \cdots \cap S_{g^p}$ peut être paramétré de manière C^1 à l'aide de $n - p$ variables réelles. En déduire que l'ensemble des vecteurs dérivés en $t = 0$ pour les arcs C^1 $t \mapsto M(t)$ tracés sur $S_{g^1} \cap S_{g^2} \cdots \cap S_{g^p}$ vérifiant $M(0) = M_0$ est un espace vectoriel de dimension $n - p$, qui n'est autre que

$$\bigcap_{i=1}^{p} \ker dg^i_{M_0}$$

(intersection de p hyperplans indépendants). Si M_0 est un extremum local de la fonction $f|_{S_{g_1}\cap S_{g^2}\cdots\cap S_{g^p}}$, montrer que, pour un tel arc paramétré,

$$df_{M_0}\left(\overrightarrow{M}'(0)\right)=0$$

En déduire que

$$\bigcap_{i=1}^{p}\ker dg^i_{M_0}\subset\ker df_{M_0}$$

et en déduire qu'il existe des réels $(\lambda_i)_{1\leqslant i\leqslant p}$ (appelés "multiplicateurs de Lagrange") tels que

$$df_{M_0}=\sum_{i=1}^{p}\lambda_i dg^i_{M_0}$$

On obtient ainsi une condition nécessaire pour que M_0 soit extremum de f sous les contraintes $g_i=0$.

EXERCICE 18-3.42 (application de l'exercice précédent). Soit $\mathbb{E}_n$ un espace euclidien de dimension orienté. On définit l'application

$$f:\mathbb{E}_n^n\to\mathbb{R}\quad(x_1,\ldots,x_n)\mapsto[x_1,\ldots,x_n]\ \text{(produit mixte des vecteurs } x_i)$$

En utilisant le calcul différentiel, trouver le maximum de f sous les contraintes

$$\|x_1\|^2=\cdots=\|x_n\|^2=1$$

Retrouver ce résultat par un raisonnement géométrique (voir exercice **14-4**.2). En déduire l'inégalité dite d'Hadamard

$$\text{pour } A=(a_{ij})\in\mathcal{M}_n(\mathbb{R})\quad|\det A|^2\leqslant\prod_{j=1}^{n}\left(\sum_{i=1}^{n}a_{ij}^2\right)$$

et étudier le cas de l'égalité.

18-3.5 Exercice : caractérisation des fonctions homogènes

DÉFINITION 18-3.43 *Soit C un ouvert de $\mathbb{R}^n$ stable par les homothéties de rapports strictement positifs (on dit que C est un* ***cône ouvert****). Si $\alpha\in\mathbb{R}$, une application*

$$f:C\to\mathbb{R}\quad x=(x_1,\ldots,x_n)\mapsto f(x_1,\ldots,x_n)$$

est dite (positivement) homogène de degré α si et seulement si

$$\forall x\in C\quad\forall t>0\quad f(tx)=f(tx_1,\ldots,tx_n)=t^\alpha f(x)$$

Montrer qu'une application $f : C \to \mathbb{R}$ de classe C^1 est positivement homogène de degré α si et seulement si elle vérifie l'identité d'Euler

$$\forall x \in C \quad df_x(x) = \alpha f(x)$$

ce qui peut s'écrire également

$$\forall (x_1, \dots, x_n) \in C \quad \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = \alpha f(x_1, \dots, x_n)$$

Indication : Si $f \in C^1(C, \mathbb{R})$ est positivement homogène, dériver la fonction

$$\mathbb{R}^{*+} \ni t \mapsto f(tx_1, \dots, tx_n)$$

pour $x = (x_1, \dots, x_n)$ fixé dans C. Réciproquement, si f vérifie l'identité d'Euler, considérer l'application φ définie sur $]0, +\infty[$ par $\varphi(t) = t^{-\alpha} f(tx_1, \dots, tx_n)$, avec $(x_1, \dots, x_n) \in C$.

18-4 Exercices

EXERCICE 18-4.1 Etudier la différentiabilité de la fonction définie par

$$f(x, y) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{x + ny} \quad (x \text{ et } y \text{ réels strictement positifs})$$

EXERCICE 18-4.2 Soit φ une fonction de classe C^2 de $\mathbb{R}$ dans $\mathbb{R}$, et l'espace vectoriel $E = C^0([a, b], \mathbb{R})$ normé par la norme de la convergence uniforme. On définit $\Phi : E \to \mathbb{R}$ par

$$\Phi(f) = \int_a^b \varphi(f(t))dt$$

Etudier la différentiabilité de Φ.

EXERCICE 18-4.3 Soit $N \in \{N_\infty, N_1, N_2\}$ une des normes usuelles sur $\mathbb{R}^n$. Etudier l'ensemble des points de $\mathbb{R}^n$ où N est différentiable et expliciter dN.

EXERCICE 18-4.4 Exemples d'utilisation de changements de variables :

1. Trouver $f \in C^2(\mathbb{R}, \mathbb{R})$ telle que $F(x, y) = f(\dfrac{\cos 2x}{\text{ch } 2y})$ soit harmonique dans $\mathbb{R}^2$.
2. Même question avec $F(x_1, \dots, x_n) = f(x_1^2 + \dots + x_n^2)$ sur $\mathbb{R}^n$.
3. Résoudre $\dfrac{\partial^2 f}{\partial x^2} - 3\dfrac{\partial^2 f}{\partial x \partial y} + 2\dfrac{\partial^2 f}{\partial y^2} = x^2 + 2xy + y^2$
4. Résoudre : $x^2\dfrac{\partial^2 f}{\partial x^2} - y^2\dfrac{\partial^2 f}{\partial y^2} = x\dfrac{\partial f}{\partial x} - y\dfrac{\partial f}{\partial y}$
5. Résoudre : $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = 2\dfrac{x^3 + y^3 - x^2 y}{(x + y)^5}$ sur $\{x > 0, y > 0\}$

6. Trouver les fonctions $f \in \mathcal{C}^1(\mathbb{R}^2 - \{(0,0)\}, \mathbb{C})$ telles que $x\dfrac{\partial f}{\partial y} - y\dfrac{\partial f}{\partial x} = kf$ où $k \in \mathbb{C}$ est une constante arbitraire.

EXERCICE 18-4.5 Soit N une norme sur $\mathbb{R}^n$ telle que $x \mapsto N^2(x)$ soit de classe $\mathcal{C}^2$ sur $\mathbb{R}^n$. Montrer que N est une norme euclidienne.

EXERCICE 18-4.6 Soit U un ouvert convexe de R^n contenant 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^\infty$ telle que $f(0) = 0$, toutes les dérivées partielles de f d'ordre inférieur ou égal à $p-1$ étant nulles en 0. Montrer qu'il existe des applications $u_\alpha \in \mathcal{C}^\infty(U, \mathbb{R})$ telles que

$$\forall x \in U \quad f(x) = \sum_{|\alpha|=p} u_\alpha(x) x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

où $x = (x_1, \ldots, x_n)$ et $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. On note $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

EXERCICE 18-4.7 Soit f de classe $\mathcal{C}^2$ d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$. On suppose que f présente un minimum local en $(x_0, y_0) \in U$. Montrer que $\Delta f(x_0, y_0) \geqslant 0$. Soit $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$ et f continue : $D \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur l'intérieur de D et harmonique dans ce domaine. Montrer que, si f est nulle sur la frontière de D, alors f est identiquement nulle sur D. (On pourra considérer, pour $\varepsilon > 0$, la fonction $g(x,y) = f(x,y) + \varepsilon(x^2 + y^2 - 1)$ et utiliser le résultat précédent).

EXERCICE 18-4.8 Soit $n \in \mathbb{N}$.

- Trouver $g \in \mathcal{C}^2(]0,\pi[, \mathbb{C})$ avec $\sin\theta \dfrac{d}{d\theta}(\sin\theta \dfrac{dg}{d\theta}) = n^2 g(\theta)$.
- Soit $f \in \mathcal{C}^2(\mathbb{R}\times]0,\pi[, \mathbb{C})$ telle que

$$\forall \theta \in]0,\pi[\quad f(\bullet, \theta) \text{ est } 2\pi\text{-périodique}$$

$$\frac{\partial^2 f}{\partial \varphi^2} + \sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial f}{\partial\theta}\right) = 0 \text{ en tout point } (\varphi, \theta) \in \mathbb{R}\times]0,\pi[$$

  Montrer que, si on suppose f bornée, alors f est constante.

EXERCICE 18-4.9 Soit $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$ vérifiant $\forall t \in \mathbb{R} \quad |f'(t)| \leqslant k < 1$. On définit $F : \mathbb{R}^2 \to \mathbb{R}^2$ par $F(x,y) = (x - f(y), y - f(x))$. Montrer que F est un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

EXERCICE 18-4.10 Soit U un ouvert convexe d'un evn E et $f : U \to \mathbb{R}$ convexe. On suppose que f est différentiable en un point $a \in U$, où la différentielle df_a est nulle. Montrer que f présente un minimum absolu en a.

EXERCICE 18-4.11 Soit $\varphi : \mathbb{R} \to \mathcal{M}_n(\mathbb{C})$ de classe $\mathcal{C}^k$ telle que, pour tout $t \in \mathbb{R}$ $\varphi(t)$ ait n valeurs propres simples. Montrer que tout $t_0 \in \mathbb{R}$ possède un voisinage sur lequel on peut définir une fonction $t \mapsto (X_1(t), \ldots, X_n(t))$ de classe $\mathcal{C}^k$ définissant une base de vecteurs propres de $\varphi(t)$. (Considérer le polynôme caractéristique de $\varphi(t)$ et appliquer le théorème des fonctions implicites).

EXERCICE 18-4.12 Sur le cercle unité du plan euclidien rapporté à un repère orthonormé on considère trois points A, B et C. A quelle condition le périmètre du triangle ABC est-il maximum ? Même question avec quatre points.

EXERCICE 18-4.13 Dans un plan affine euclidien, on considère un triangle d'intérieur U et de côtés $\mathcal{D}_1, \mathcal{D}_2$ et $\mathcal{D}_3$. Pour $M \in U$ on pose

$$f(M) = d(M, \mathcal{D}_1)d(M, \mathcal{D}_2)d(M, \mathcal{D}_3)$$

Montrer que dans un repère du plan f est polynomiale et atteint sur U un maximum. Montrer que ce maximum est unique et indiquer le point où il est atteint.

EXERCICE 18-4.14 Si A, B, C sont les sommets d'un triangle du plan euclidien $\mathcal{P}$, pour $M \in \mathcal{P}$ on pose

$$g(M) = MA + MB + MC$$

1. Montrer que f admet un minimum, mais qu'il ne peut être atteint en un point extérieur au triangle.
2. Montrer que, si les trois angles du triangle ont une mesure $< \dfrac{2\pi}{3}$, il existe dans l'intérieur de U un unique point où ce minimum est atteint.
3. Si l'un des angles est de mesure $\geqslant \dfrac{2\pi}{3}$, vérifier que c'est au sommet de cet angle que le minimum est atteint.

EXERCICE 18-4.15 Si les côtés d'un triangle ont pour longueur a, b et c et sa surface vaut S, montrer que

$$S \leqslant \frac{\sqrt{3}}{12}(a^2 + b^2 + c^2)$$

$$S \leqslant \frac{\sqrt{3}}{4}(abc)^{\frac{2}{3}}$$

avec égalité ssi le triangle est équilatéral.

EXERCICE 18-4.16 Les a_i sont des réels > 0 avec $\sum_{i=1}^{n} a_i = 1$. Montrer que

$$\prod_{i=1}^{n} a_i(1 - a_i) \leqslant \frac{(n-1)^n}{n^{2n}}$$

EXERCICE 18-4.17 Montrer que toute fonction de classe $\mathcal{C}^\infty$ sur $(\mathbb{R}^{*+})^2$ vérifiant

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = f$$

est somme de deux fonctions homogènes de degrés à préciser.

EXERCICE 18-4.18 On note E l'ensemble des fonctions f vérifiant les conditions :

$$\begin{cases} \text{(i)} & f \text{ est continue } : [0, \pi] \times [0, +\infty[\longrightarrow \mathbb{R} \\ \text{(ii)} & f \text{ est } \mathcal{C}^2 \text{ sur }]0, \pi[\times]0, +\infty[\\ \text{(iii)} & \forall\, (x, t) \in]0, \pi[\times]0, +\infty[\quad \dfrac{\partial^2 f}{\partial x^2}(x, t) = \dfrac{\partial f}{\partial t}(x, t) \\ \text{(iv)} & \text{Pour } t \geqslant 0 \quad f(0, t) = f(\pi, t) = 0 \\ \text{(v)} & \forall\, x \in [0, \pi] \quad \lim\limits_{t \to +\infty} f(x, t) = 0 \end{cases}$$

1. Quels sont les éléments non nuls de E de la forme $(x, t) \mapsto g(x)h(t)$?

2. Développer en série de Fourier la fonction φ impaire, 2π-périodique qui coïncide avec $x \mapsto x(\pi - x)$ sur $[0, \pi]$.

3. En déduire une fonction $f \in E$ telle que $\forall x \in [0, \pi]$ on ait $f(x, 0) = x(\pi - x)$.

EXERCICE 18-4.19 Soit f de classe C^1 de $\mathbb{R}^n$ dans lui même vérifiant

$$\forall x, y \in \mathbb{R}^n \qquad \|f(x) - f(y)\| \geqslant \|x - y\|$$

Montrer que f est un C^1-difféomorphisme de $\mathbb{R}^n$ dans lui-même.

EXERCICE 18-4.20 Montrer que pour $|t| < \dfrac{1}{\sqrt{2}}$, l'équation

$$\sin(tx) + \cos(tx) = x$$

admet une unique solution $x = \varphi(t)$. Montrer que φ est C^∞ et en donner un développement limité à l'ordre 3 en 0.

EXERCICE 18-4.21 Trouver les fonctions de classe C^2 de $(\mathbb{R}^{*+})^2$ dans $\mathbb{R}$ vérifiant

$$x^2 \frac{\partial^2 f}{\partial x^2} - y^2 \frac{\partial^2 f}{\partial y^2} = 0$$

(On pourra utiliser le changement de variables $xy = u$; $\dfrac{y}{x} = v$).

Chapitre 19

Equations différentielles

19-1 Equations linéaires

19-1.1 Etude générale

19-1.1.1 Définition

Dans toute cette section, $(\mathbb{E}, \| \|)$ est un espace normé complet. Ce sera souvent $\mathbb{R}$ ou $\mathbb{C}$ (les équations considérées seront alors dites "scalaires"), ou un espace de dimension finie.

DÉFINITION 19-1.1 *On appelle équation différentielle linéaire du premier ordre toute équation de la forme*

$$(E) \quad x'(t) = a(t)(x(t)) + b(t)$$

(qu'on notera souvent en abrégé $x' = ax + b$) où

$$\begin{array}{llll} a: & I \to \mathcal{L}_c(\mathbb{E}) & t \mapsto a(t) \\ b: & I \to \mathbb{E} & t \mapsto b(t) \end{array}$$

sont deux fonctions continues définies sur un intervalle de $\mathbb{R}$ (non réduit à un point), l'espace $\mathcal{L}_c(\mathbb{E})$ étant muni de la norme[1] *des endomorphismes continus sur $\mathbb{E}$.*

[1]Ou de toute autre norme si $\mathbb{E}$ est de dimension finie.

DÉFINITION 19-1.2 *Une solution de l'équation (E) est une application dérivable $y : I \to \mathbb{E}$ vérifiant*

$$\forall t \in I \quad y'(t) = a(t)(y(t)) + b(t)$$

Pour simplifier un peu les notations, on écrira dans ce qui suit $a(t)y(t)$ le vecteur de $\mathbb{E}$ image de $y(t)$ par l'endomorphisme $a(t)$.

Dans le cas (le plus fréquent pour nous) où l'espace $\mathbb{E}$ est de dimension p, et est rapporté à une base $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant p}$, l'endomorphisme $a(t)$ peut être représenté par sa matrice $A(t) \in \mathcal{M}_p(\mathbb{K})$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

$$A(t) = (a_{ij}(t))_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant p}}$$

et la continuité de l'application $t \mapsto a(t) \in \mathcal{L}_c(\mathbb{E}) = \mathcal{L}(\mathbb{E})$ se traduit évidemment par la continuité des p^2 applications $t \mapsto a_{ij}(t)$ de I dans $\mathbb{K}$. De même, le vecteur $b(t)$ est repéré par la colonne de ses coordonnées dans $\mathcal{B}$

$$B(t) = \begin{pmatrix} b_1(t) \\ \vdots \\ b_p(t) \end{pmatrix} \in \mathbb{K}^p$$

soit p applications continues $I \ni t \mapsto b_i(t) \in \mathbb{K}$. L'équation (E) s'écrit alors, si

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_p(t) \end{pmatrix}$$

est la colonne des coordonnées (fonctions dérivables de la variable t) du vecteur "inconnu" $x(t)$ dans $\mathcal{B}$,

$$X'(t) = A(t)X(t) + B(t)$$
$$\Leftrightarrow \left\{ \begin{array}{c} x_1'(t) = a_{11}(t)x_1(t) + \cdots + a_{1j}(t)x_j(t) + \cdots + a_{1p}(t)x_p(t) + b_1(t) \\ \vdots \\ x_i'(t) = a_{i1}(t)x_1(t) + \cdots + a_{ij}(t)x_j(t) + \cdots + a_{ip}(t)x_p(t) + b_i(t) \\ \vdots \\ x_p'(t) = a_{p1}(t)x_1(t) + \cdots + a_{pj}(t)x_j(t) + \cdots + a_{pp}(t)x_p(t) + b_p(t) \end{array} \right.$$

ce qui se présente donc comme un système de p équations différentielles scalaires du premier ordre, à p fonctions inconnues $(t \mapsto x_i(t))_{1 \leqslant i \leqslant p}$. Le cas $p = 1$ est évidemment important, et l'équation s'écrit alors

$$x'(t) = a(t)x(t) + b(t) \quad \text{avec} \quad a, b \in \mathcal{C}^0(I, \mathbb{K})$$

PROPOSITION 19-1.3 Toute fonction $I \ni t \mapsto y(t)$ solution de (E) est dans $\mathcal{C}^1(I, \mathbb{E})$.

Démonstration : Par hypothèse y est dérivable sur I et

$$\forall t \in I \quad y'(t) = a(t)y(t) + b(t)$$

y' est donc une fonction continue sur I (l'application $t \mapsto a(t)\, y(t)$ est continue, comme composée de l'application $I \to \mathcal{L}_c(\mathbb{E}) \times \mathbb{E}$ définie par $t \mapsto (a(t), y(t))$ et de l'application bilinéaire continue $\mathcal{L}_c(\mathbb{E}) \times \mathbf{E} \to \mathbb{E}$ définie par $(u, z) \mapsto uz = u(z)$). ■

De même, si $a \in \mathcal{C}^p(I, \mathcal{L}_c(\mathbb{E}))$ et $b \in \mathcal{C}^p(I, \mathbb{E})$, toute solution de (E) est de classe $\mathcal{C}^{p+1}$ sur I.

19-1.1.2 Théorème de Cauchy

Par un procédé d'approximations successives, nous allons prouver l'existence et l'unicité d'une solution de (E) pour une "condition initiale" donnée. Pour les besoins de la démonstration, nous allons d'abord transformer (E) en une équation intégrale[2] équivalente (essentiellement parce qu'une opération d'intégration sur un segment se comporte bien par rapport à la norme de la convergence uniforme, ce qui n'est pas le cas pour la dérivation).

LEMME 19-1.4 On considère l'équation linéaire

$$(E) \quad x' = ax + b$$

sur un intervalle I de $\mathbb{R}$ et soient $(t_0, y_0) \in I \times \mathbb{E}$. Une application $y : I \to \mathbf{E}$ vérifiant $y(t_0) = y_0$ est solution de (E) si et seulement si y est continue sur I et si

$$(E_I) \quad \forall t \in I \quad y(t) = y_0 + \int_{t_0}^{t} [a(u)\, y(u) + b(u)]\, du$$

Démonstration : Si y est solution de (E) elle est de classe C^1 sur I et vérifie donc

$$\forall t \in I \quad y(t) - y(t_0) = \int_{t_0}^{t} y'(u)\, du = \int_{t_0}^{t} [a(u)\, y(u) + b(u)]\, du$$

ce qui prouve la nécessité de (E_I). Réciproquement, si $y \in \mathcal{C}^0(I, \mathbf{E})$, l'application définie par $u \mapsto a(u)\, y(u) + b(u)$ est continue sur I, donc intégrable sur tout segment inclus dans I. Si y vérifie (E_I), on a évidemment $y(t_0) = y_0$, et $t \mapsto y(t)$ est dans $\mathcal{C}^1(I, \mathbb{E})$, comme intégrale d'une fonction continue dépendant d'une de ses bornes. On a alors

$$\forall t \in I \quad y'(t) = a(t)\, y(t) + b(t)$$

ce qui montre que y est solution de (E). ■

THÉORÈME 19-1.5 Si $a \in \mathcal{C}^0(I, \mathcal{L}_c(\mathbb{E}))$ et $b \in \mathcal{C}^0(I, \mathbb{E})$. pour tout couple $(t_0, y_0) \in I \times \mathbb{E}$, l'équation différentielle

$$(E) \quad x'(t) = a(t)\, x(t) + b(t)$$

possède une seule solution $y \in \mathcal{C}^1(I, \mathbb{E})$ vérifiant la condition initiale

$$y(t_0) = y_0$$

[2] On utilise donc le fait que $(E, \| \|)$ est complet.

Démonstration : On montre que l'équation (E_I) possède une unique solution $y \in \mathcal{C}^0(I, \mathbb{E})$. Pour prouver l'existence, nous utilisons un procédé itératif, en construisant une suite de fonctions continues sur I, qui va converger uniformément sur tout segment inclus dans I vers une fonction solution de (E_I). Pour cela, définissons la fonction $z_0 : I \to \mathbb{E}$ par

$$\forall t \in I \quad z_0(t) = y_0$$

Cette fonction est évidemment continue sur I. Pour $n \in \mathbb{N}$, si on suppose connue la fonction $z_n \in \mathcal{C}^0(I, \mathbb{E})$, définissons $z_{n+1} : I \to \mathbb{E}$ par

$$\forall t \in I \quad z_{n+1}(t) = y_0 + \int_{t_0}^{t} [a(u)\, z_n(u) + b(u)]\, du$$

(intégrale d'une fonction continue sur un segment). La fonction z_{n+1} est alors évidemment continue (et même de classe $\mathcal{C}^1$) sur I. On a ainsi défini, par récurrence, une suite de fonctions $(z_n)_{n\in\mathbb{N}}$ continues sur I. Montrons que cette suite converge simplement sur I, avec convergence uniforme sur tout segment inclus dans I. Pour cela, considérons la série télescopique associée, de terme général

$$w_n(t) = z_{n+1}(t) - z_n(t)$$

et montrons que cette série converge normalement sur tout segment inclus dans I. Si K est un segment inclus dans I, considérons K_0 le plus petit segment inclus dans I contenant à la fois K et t_0. Comme $t \mapsto a(t)$ est continue, cette fonction est bornée sur K_0 ($\mathcal{L}_c(\mathbb{E})$ est muni de la norme des endomorphismes continus de $(E, \| \|)$) et

$$\exists M > 0 \quad \forall u \in K_0 \quad \|a(u)\| \leqslant M$$

Pour $t \in I$, on a

$$\begin{aligned} w_{n+1}(t) &= z_{n+2}(t) - z_{n+1}(t) \\ &= \int_{t_0}^{t} a(u)\,(z_{n+1}(u) - z_n(u))\, du = \int_{t_0}^{t} a(u)\, w_n(u)\, du \end{aligned}$$

d'après la relation de récurrence définissant la suite $(z_n)_{n\in\mathbb{N}}$. La fonction w_0 étant continue sur I, on peut également trouver un réel $N > 0$ tel que

$$\forall u \in K_0 \quad \|w_0(u)\| \leqslant N$$

On a alors

$$\begin{aligned} \forall t \in K_0 \quad \|w_1(t)\| &\leqslant \left| \int_{t_0}^{t} \|a(u)\, w_0(u)\|\, du \right| \\ &\leqslant \left| \int_{t_0}^{t} \|a(u)\|\, \|w_0(u)\|\, du \right| \leqslant MN\,|t - t_0| \end{aligned}$$

puisque le segment d'intégration est inclus dans K_0 avec

$$\forall u \in K_0 \quad \|a(u)\, w_0(u)\| \leqslant \|a(u)\|\, \|w_0(u)\| \leqslant MN$$

On en déduit

$$\forall t \in K_0 \quad \|w_2(t)\| \leqslant \left| \int_{t_0}^{t} \|a(u)\|\, \|w_1(u)\|\, du \right| \leqslant \left| \int_{t_0}^{t} M^2 N\, |u - t_0|\, du \right| = N \frac{M^2 |t - t_0|^2}{2!}$$

et, par récurrence sur n, on obtient

$$\forall t \in K_0 \quad \|w_n(t)\| \leqslant N \frac{M^n |t - t_0|^n}{n!} \leqslant N \frac{(ML)^n}{n!}$$

où L est la longueur du segment K_0. Cette inégalité prouve la convergence normale sur K_0 (et donc sur K) de la série de fonctions de terme général $w_n(t)$, donc la convergence uniforme de la suite de fonctions z_n, puisque

$$\forall t \in I \quad z_n(t) = y_0 + \sum_{k=0}^{n-1} w_k(t)$$

La fonction y définie sur I par

$$y(t) = \lim_{n \to +\infty} z_n(t)$$

est donc continue sur I (limite d'une suite de fonctions continues avec convergence uniforme locale). De plus, si $K \subset I$ est un segment, l'inégalité

$$\forall u \in K \quad \|a(u)\, z_n(u) - a(u)\, y(u)\| \leqslant \|a(u)\|\, \|z_n(u) - y(u)\| \leqslant \sup_K \|a(u)\| \times \|z_n - y\|_{\infty,K}$$

montre que la suite $(az_n + b)_{n \in \mathbb{N}}$ converge uniformément sur K vers la fonction $(ay + b)$. Si $t \in K$, en passant à la limite dans l'égalité

$$\forall n \in \mathbb{N} \quad z_{n+1}(t) = y_0 + \int_{t_0}^{t} [a(u)\, z_n(u) + b(u)]\, du$$

on obtient

$$y(t) = y_0 + \int_{t_0}^{t} [a(u)\, y(u) + b(u)]\, du$$

ce qui montre que y vérifie (E_I). L'unicité se prouve par une technique analogue : si y et z sont deux solutions de (E_I), la fonction $\delta = y - z$ vérifie

$$\forall t \in I \quad \delta(t) = \int_{t_0}^{t} a(u)\, \delta(u)\, du$$

Avec les mêmes notations que précédemment, si on note

$$N_0 = \sup_{u \in K_0} \|\delta(u)\|$$

(qui existe puisque δ est continue sur le segment K_0), on montre par récurrence que

$$\forall n \in \mathbb{N} \quad \forall t \in K_0 \quad \|\delta(t)\| \leqslant N \frac{(M|t - t_0|)^n}{n!}$$

ce qui donne $\delta(t) = 0_{\mathbb{E}}$ pour tout t, et montre que $y = z$. ■

REMARQUE 19-1.6 Si J est un sous-intervalle de I (non réduit à un point), le théorème de Cauchy montre qu'une fonction z de classe $\mathcal{C}^1$ de J dans $\mathbb{E}$ qui vérifie (E) sur J, c'est à dire

$$\forall t \in J \quad z'(t) = a(t)\,z(t) + b(t)$$

se prolonge de manière unique en une solution y sur I : il suffit en effet de fixer $t_0 \in J$ et de prendre $y_0 = z(t_0)$. Le théorème précédent assure l'existence d'une solution y sur I vérifiant cette condition initiale. Sa restriction à J est évidemment solution de (E) sur J, et vérifiant la même condition initiale que z est égale à z (théorème de Cauchy appliqué à l'équation sur J). On a donc $y|_J = z$ et y prolonge bien z.

19-1.1.3 Equation homogène : espace des solutions

Pourquoi l'équation (E) est-elle qualifiée de linéaire ? C'est parce qu'on peut l'interpréter en termes de morphisme et d'espaces vectoriels : si les fonctions $a \in \mathcal{C}^0(I, \mathcal{L}_c(\mathbb{E}))$ et $b \in \mathcal{C}^0(I, \mathbb{E})$ sont données, l'application

$$\Phi : \mathcal{C}^1(I, \mathbb{E}) \to \mathcal{C}^0(I, \mathbb{E}) \quad x \mapsto z = \Phi(x)$$

définie par

$$\forall t \in I \quad z(t) = x'(t) - a(t)\,x(t)$$

est évidemment linéaire, et résoudre (E), c'est résoudre l'équation linéaire

$$\Phi(x) = b$$

Le théorème de Cauchy montre en particulier la surjectivité de Φ. L'ensemble des solutions de (E) est donc un sous-espace affine de $\mathcal{C}^1(I, \mathbb{E})$, dont la direction est le noyau de Φ, c'est à dire l'ensemble des solutions de l' **équation homogène associée à** (E), dite aussi équation "sans second membre"

$$(E_H) : \Phi(x) = 0 \Leftrightarrow x'(t) = a(t)\,x(t)$$

Si on connaît une solution x_1 "particulière" de (E), toute autre solution s'en déduira facilement si on connaît les solutions de (E_H), puisque

$$x \text{ solution de } (E) \Leftrightarrow x \in x_1 + \ker \Phi$$

On s'intéresse donc dans cette section à la structure de l'ensemble $\mathcal{S}_H$ des solutions de (E_H).

THÉORÈME 19-1.7 L'ensemble des solutions de l'équation homogène est un sous-espace vectoriel de $\mathcal{C}^1(I,\mathbb{E})$ isomorphe à $\mathbb{E}$: si t_0 est un point arbitrairement fixé dans I, la fonction d'évaluation en t_0

$$\varphi_{t_0} : \mathcal{S}_H \to \mathbb{E} \quad x \mapsto x(t_0)$$

est un isomorphisme d'espaces vectoriels. En particulier, si l'espace $\mathbb{E}$ est de dimension finie

$$\dim \mathcal{S}_H = \dim \mathbb{E}$$

Démonstration : Le théorème de Cauchy appliqué à l'équation (E_H) montre que, pour tout $x_0 \in \mathbb{E}$, il existe un seul élément x de $\mathcal{S}_H$ vérifiant $x(t_0) = x_0$. L'application φ_{t_0} (évidemment linéaire) est donc un isomorphisme de $\mathcal{S}_H = \ker \Phi$ dans $\mathbb{E}$. ■

REMARQUE 19-1.8 Cet isomorphisme est souvent un moyen commode de décrire l'évolution "temporelle" d'une solution $t \mapsto x(t)$ de (E_H) en ayant une vision "photographique instantanée" de la trajectoire (voire "stroboscopique" si des phénomènes périodiques sont en jeu) : si une solution de (E_H) passe à l'instant $t_0 \in I$ par la position x_0, quelle sera sa position x_1 à un instant $t_1 > t_0$ (ou quelle était sa position à cet instant si $t_1 < t_0$) ? Le théorème précédent montre que la correspondance

$$R(t_0, t_1) : \mathbb{E} \to \mathbb{E} \quad x_0 \mapsto x_1$$

est d'abord parfaitement définie[3], et est un isomorphisme d'espace vectoriel. Avec les notations précédentes

$$R(t_0, t_1) = \varphi_{t_1} \circ \varphi_{t_0}^{-1}$$

La définition même de cette correspondance (ou, de manière plus algébrique, l'égalité précédente) montre que, pour t_0, t_1 et $t_2 \in I$,

$$R(t_1, t_0) = [R(t_0, t_1)]^{-1} \quad \text{et} \quad R(t_0, t_2) = R(t_1, t_2) \circ R(t_0, t_1)$$

Une conséquence fondamentale du théorème précédent est que l'indépendance d'une famille de solutions de (E_H) peut se lire à un instant t_0 arbitraire :

COROLLAIRE 19-1.9 Si $x_1, \ldots, x_p$ sont p solutions de l'équation (E_H), elles sont indépendantes dans l'espace $\mathcal{S}_H$ (ou, de manière équivalente, dans $\mathcal{C}^1(I,\mathbb{E})$) si et seulement s'il existe t_0 dans I tel que

$$(x_i(t_0))_{1 \leqslant i \leqslant p} \text{ est une famille libre de } \mathbb{E}$$

Si t_0 existe tel que $(x_i(t_0))_{1 \leqslant i \leqslant p}$ soient indépendants, tous les points de I vérifient cette propriété.

[3]En d'autres termes, un système (mécanique par exemple) régi par une équation différentielle linéaire est déterministe : son état à l'instant t_0 détermine entièrement son évolution ultérieure.

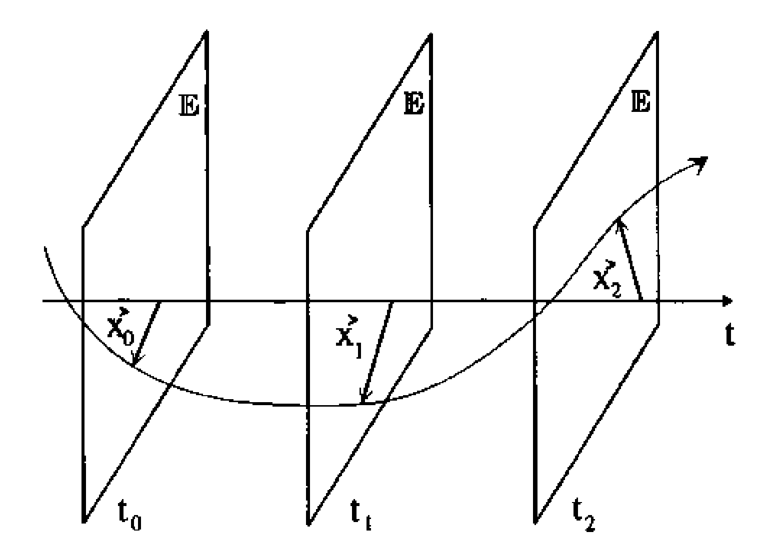


Figure **19.1** – *Evolution d'une solution de $x' = ax$*

Démonstration : Si t_0 existe, l'indépendance des solutions $(x_i)_{1\leqslant i\leqslant p}$ en découle immédiatement, puisqu'une relation de liaison entre elles donnerait

$$\forall t \in I \quad \sum_{i=1}^{p} \lambda_i x_i(t) = 0_{\mathbb{E}} \Rightarrow \sum_{i=1}^{p} \lambda_i x_i(t_0) = 0_{\mathbb{E}} \Rightarrow \forall i \quad \lambda_i = 0$$

Le corollaire trouve donc son intérêt dans l'implication réciproque : si $(x_i)_{1\leqslant i\leqslant p}$ est une famille libre, alors pour tout t_0 de I la famille $\left(\varphi_{t_0}(x_i)\right)_{1\leqslant i\leqslant p}$ est libre, comme image d'une famille libre par un isomorphisme d'espace vectoriel. ■

De même, **si** $(x_i)_{1\leqslant i\leqslant p}$ **est une famille de solutions de** (E_H) **vérifiant à un instant** t_0 **une relation de liaison**

$$\sum_{i=1}^{p} \lambda_i x_i(t_0) = 0_{\mathbb{E}}$$

on aura, pour tout t **dans** I

$$\sum_{i=1}^{p} \lambda_i x_i(t) = 0_{\mathbb{E}}$$

puisque la fonction $x = \sum_{i=1}^{p} \lambda_i x_i$ est solution de (E_H) vérifiant la condition initiale $x(t_0) = 0_{\mathbb{E}}$ et est donc la solution identiquement nulle (unicité de la solution d'après le théorème de Cauchy).

Ainsi, les fonctions x_1 et $x_2 \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}^2)$ définies par

$$\forall t \in \mathbb{R} \quad x_1(t) = (1, t) \text{ et } x_2(t) = (2, 2t^3)$$

sont clairement indépendantes, mais ne peuvent être solutions d'une même équation différentielle linéaire **homogène**, puisque les vecteurs $x_1(1)$ et $x_2(1)$ sont liés.

19-1.1.4 Dimension finie : système fondamental de solutions

Lorsque $\mathbb{E} = \mathbb{E}_n$ est de dimension finie n (ce que nous supposerons dans cette section), l'espace vectoriel $\mathcal{S}_H$ est également de dimension n.

DÉFINITION 19-1.10 *On appelle système fondamental de solutions de l'équation homogène (E_H) toute base de l'espace vectoriel des solutions $\mathcal{S}_H$.*

Si $(x_i)_{1\leqslant i\leqslant n}$ est un système fondamental de solutions, toute solution $x \in \mathcal{S}_H$ s'écrit (de manière unique) comme combinaison linéaire de $(x_i)_{1\leqslant i\leqslant n}$:

$$\exists\ (\lambda_i)_{1\leqslant i\leqslant n} \in \mathbb{K}^n \quad \forall t \in I \quad x(t) = \sum_{i=1}^{n} \lambda_i x_i(t)$$

Le corollaire **19-1**.9 montre qu'une famille $(x_i)_{1\leqslant i\leqslant n}$ de solutions de (E_H) est un système fondamental de solutions si et seulement s'il existe $t_0 \in I$ tel que

$$\mathcal{B}_{t_0} = (x_i(t_0))_{1\leqslant i\leqslant n} \text{ est une base de } \mathbb{E}_n$$

Lorsque cette propriété est vérifiée, on a

$$\forall t \in I \quad \mathcal{B}_t = (x_i(t))_{1\leqslant i\leqslant n} \text{ est une base de } \mathbb{E}_n$$

et une fonction $I \ni t \mapsto x(t) \in \mathbb{E}_n$ est solution de (E_H) si et seulement si les cordonnées du vecteur ("variable") $x(t)$ sont constantes dans la base ("mobile") $\mathcal{B}_t$.

DÉFINITION 19-1.11 *Soit $\mathcal{B}$ une base fixée de $\mathbb{E}_n$. Si $(x_i)_{1\leqslant i\leqslant n}$ sont n solutions de (E_H), on appelle wronskien de ces solutions dans la base $\mathcal{B}$ l'application*

$$W : I \to \mathbb{K} \quad t \mapsto W(t) = \det_{\mathcal{B}}(x_1(t), \ldots, x_n(t))$$

On a alors deux possibilités qui s'excluent mutuellement :

- Il existe $t_0 \in I$ avec $W(t_0) \neq 0$. Les $(x_i)_{1\leqslant i\leqslant n}$ forment alors un système fondamental de solutions de (E_H), et pour tout $t \in I$ le déterminant wronskien $W(t)$ est non nul.
- Il existe $t_0 \in I$ avec $W(t_0) = 0$. Les vecteurs $(x_i(t_0))_{1\leqslant i\leqslant n}$ vérifient alors une relation de liaison non triviale, qui est aussi vérifiée pour tout t par la famille $(x_i(t))_{1\leqslant i\leqslant n}$. On a donc $\forall t \in I \quad W(t) = 0$.

PROPOSITION 19-1.12 (Formule de Liouville) Si $(x_i)_{1\leqslant i\leqslant n}$ est une famille de solutions de l'équation homogène

$$(E_H) \quad x'(t) = a(t)\,x(t)$$

et si W est leur wronskien dans une base de $\mathbb{E}_n$, on a

$$\forall t, t_0 \in I \quad W(t) = W(t_0) \exp\left(\int_{t_0}^{t} \operatorname{trace}(a(u))\ du\right)$$

Démonstration : Si $(x_i)_{1\leqslant i\leqslant n}$ est liée, la fonction W est identiquement nulle, et l'égalité à démontrer est évidente. On suppose donc que $(x_i)_{1\leqslant i\leqslant n}$ est un système fondamental de solutions, et donc

$$\forall t \in I \quad \mathcal{B}_t = (x_i(t))_{1\leqslant i\leqslant n} \text{ est une base de } \mathbb{E}_n$$

La fonction W définie par

$$\forall t \in I \quad W(t) = \det_{\mathcal{B}}(x_1(t), \dots, x_n(t))$$

est clairement de classe C^1 sur I, de dérivée

$$W'(t) = \sum_{i=1}^{n} \det_{\mathcal{B}}(x_1(t), \dots, x_i'(t), \dots, x_n(t)) = \sum_{i=1}^{n} \det_{\mathcal{B}}(x_1(t), \dots, a(t)x_i(t), \dots, x_n(t))$$

La formule de changement de bases pour les déterminants donne alors

$$W'(t) = \sum_{i=1}^{n} \det_{\mathcal{B}_t}(x_1(t), \dots, a(t)x_i(t), \dots, x_n(t)) \times \det_{\mathcal{B}}(\mathcal{B}_t)$$

soit finalement

$$W'(t) = \left(\sum_{i=1}^{n} \det_{\mathcal{B}_t}(x_1(t), \dots, a(t)x_i(t), \dots, x_n(t))\right) W(t)$$

Si $(\alpha_{ij}(t))$ est la matrice de $a(t)$ dans la base $\mathcal{B}_t$, on a

$$a(t)x_i(t) = \sum_{k=1}^{n} \alpha_{ki}(t)x_k(t)$$

ce qui donne facilement

$$\det_{\mathcal{B}_t}(x_1(t), \dots, a(t)x_i(t), \dots, x_n(t)) = \alpha_{ii}(t)$$

et finalement

$$W'(t) = \left(\sum_{i=1}^{n} \alpha_{ii}(t)\right) W(t) = (\operatorname{trace} a(t))\, W(t)$$

En dérivant la fonction

$$t \mapsto W(t) \exp\left(-\int_{t_0}^{t} \operatorname{trace}(a(u))\, du\right)$$

on vérifie aisément qu'elle est constante sur I, et comme elle vaut $W(t_0)$ en t_0, on obtient la formule de Liouville. ■

REMARQUE 19-1.13 Avec l'interprétation géométrique du déterminant évoquée dans le chapitre sur les espaces euclidiens, $W(t)$ est considéré comme le volume de l'hyper-parallélépipède construit sur les vecteurs de $\mathcal{B}_t$, et on voit comment il évolue avec t. En particulier, si la trace de a est négative sur I, ce volume est une fonction décroissante. On parlera de "contraction des volumes" au long des trajectoires des solutions de (E_H).

19-1.1.5 Dimension 1 : équations scalaires

On étudie ici le cas où l'espace $\mathbb{E}_n$ est égal au corps $\mathbb{K}$. L'équation (E) s'écrit alors

$$(E) : x' = ax + b \quad \text{avec} \quad a, b \in \mathcal{C}^0(I, \mathbb{K})$$

Dans ce cas, il est possible d'expliciter les solutions de (E_H) et de (E) à l'aide d'intégrales :

- **Résolution de l'équation sans second membre** :

La méthode a été vue pour le calcul du wronskien dans la section précédente. Si $x \in \mathcal{C}^1(I, \mathbb{K})$ et $t_0 \in I$, on considère la fonction y définie par

$$\forall t \in I \quad y(t) = x(t) \exp\left(-\int_{t_0}^{t} a(u)\, du\right)$$

On a alors facilement

$$x \text{ solution de } (E_H) \Leftrightarrow \forall t \in I \quad y'(t) = 0$$

PROPOSITION 19-1.14 Si t_0 est un point quelconque de I, les solutions de l'équation $(E_H) : x' = ax$ avec $a \in \mathcal{C}^0(I, \mathbb{K})$ sont de la forme

$$x(t) = C \exp\left(\int_{t_0}^{t} a(u)\, du\right)$$

avec $C \in \mathbb{K}$ constante arbitraire.

REMARQUE 19-1.15 On retrouve ainsi le fait que S_H est un $\mathbb{K}$-ev de dimension 1. Une utilisation précise du théorème de Cauchy permet d'ailleurs de justifier la méthode de "séparation des variables" qu'on peut utiliser pour retrouver la formule précédente lorsque $\mathbb{K} = \mathbb{R}$. Sans justification, la suite d'équivalences

$$x' = ax \Leftrightarrow \frac{x'}{x} = a \Leftrightarrow \ln \frac{x(t)}{x(t_0)} = \int_{t_0}^{t} a(u)\, du \Leftrightarrow x(t) = x(t_0) \exp\left(\int_{t_0}^{t} a(u)\, du\right)$$

n'est pas correcte (on divise par une fonction qui peut s'annuler, on prend le logarithme d'une fonction dont on n'a pas contrôlé le signe) mais a l'avantage de donner un résultat correct. On pourra justifier cette démarche en faisant précéder le "raisonnement" précédent par un argument du type :

1) Une solution de (E_H) qui s'annule en un point est la solution identiquement nulle, d'après le théorème de Cauchy.

2) Si x est une solution non identiquement nulle, $t \mapsto \dfrac{x(t)}{x(t_0)}$ est strictement positive, d'après le théorème des valeurs intermédiaires.

EXERCICE 19-1.16 Résoudre sur $\mathbb{R}$ l'équation différentielle

$$(t + i)\, x'(t) = x(t)$$

• **Résolution de l'équation avec second membre :**

C'est la méthode de variation de la constante, dont nous verrons la généralisation dans la section qui suit. On utilise le même changement de fonction inconnue que pour l'équation homogène, en cherchant la solution de (E) sous la forme

$$x(t) = C(t)\exp\left(\int_{t_0}^{t} a(u)\,du\right)$$

avec $C \in C^1(I, \mathbb{K})$, puisque toute fonction $x \in C^1(I, \mathbb{K})$ peut s'écrire sous cette forme, la fonction $C(t)$ étant simplement définie par

$$C(t) = x(t)\exp\left(-\int_{t_0}^{t} a(u)\,du\right)$$

En reportant cette expression de x dans l'équation (E), on obtient

$$\forall t \in I \quad C'(t) = b(t)\exp\left(-\int_{t_0}^{t} a(u)\,du\right)$$

ce qui donne

$$\forall t \in I \quad C(t) = C(t_0) + \int_{t_0}^{t}\left(b(u)\exp\left(-\int_{t_0}^{u} a(v)\,dv\right)\right)du$$

et permet d'obtenir facilement :

THÉORÈME 19-1.17 Soient I un intervalle de $\mathbb{R}$, $a, b \in C^0(I, \mathbb{K})$ et un couple $(t_0, x_0) \in I \times \mathbb{K}$. L'unique solution de

$$(E) \quad x' = ax + b$$

vérifiant la condition initiale $x(t_0) = x_0$ est donnée par

$$x(t) = \left(x_0 + \int_{t_0}^{t} b(u)\,e^{-A(u)}du\right)e^{A(t)}$$

où $A : I \to \mathbb{K}$ est donnée par

$$A(t) = \int_{t_0}^{t} a(v)\,dv$$

REMARQUE 19-1.18 Cas des équations "non résolues en $x'(t)$" :
Si on étudie une équation de la forme

$$\forall t \in I \quad a_1(t)\,x'(t) + b_1(t)\,x(t) + c_1(t) = 0$$

où a_1, b_1 et c_1 sont trois fonctions continues d'un intervalle I dans $\mathbb{K}$, les résultats qui précèdent s'appliqueront **en restriction à tout sous-intervalle $J \subset I$ sur lequel la fonction a_1 ne s'annule pas.** En effet, sur un tel J, l'équation peut s'écrire

$$\forall t \in J \quad x'(t) = a(t)\,x(t) + b(t)$$

où a et b sont les fonctions (continues sur J) définies par

$$a(t) = -\frac{b_1(t)}{a_1(t)} \quad \text{et} \quad b(t) = -\frac{c_1(t)}{a_1(t)}$$

Le théorème de Cauchy montre alors que, sur cet intervalle, l'ensemble des solutions de l'équation est un espace vectoriel de dimension 1. Ce résultat peut être en défaut sur l'intervalle I. Par exemple, les solutions sur $]0, +\infty[$ de l'équation

$$t\,x'(t) + x(t) = 1$$

sont les fonctions pour lesquelles la dérivée de $t \mapsto tx(t)$ est constante égale à 1. On a donc

$$x \text{ solution sur }]0, +\infty[\Leftrightarrow \exists\, C \text{ constante} \quad \forall t > 0 \quad tx(t) = t + C$$

soit

$$\forall t > 0 \quad x(t) = 1 + \frac{C}{t}$$

On retrouve bien la structure d'espace affine de dimension 1. Sur un intervalle I contenant 0, le raisonnement précédent peut être reproduit, mais on a alors nécessairement $C = 0$, ce qui montre que l'équation n'a qu'une solution sur I, et la conclusion du théorème de Cauchy n'est pas valable sur I si l'on choisit une condition initiale (t_0, x_0) avec $t_0 \in I$ et $x_0 \neq 1$.

De même, on peut vérifier que l'équation

$$tx'(t) - x(t) = 0$$

possède une infinité de solutions sur $\mathbb{R}$ qui s'annulent toutes en 0.

EXERCICE 19-1.19 Résoudre dans $]-1, 1[$, dans $]-\infty, -1[$ et $]1, +\infty[$ l'équation

$$(1 - t^2)\, x'(t) - tx(t) = 1$$

Les solutions dans $]-1, 1[$ sont de la forme

$$x(t) = \frac{\arcsin t + C_1}{\sqrt{1 - t^2}}$$

et dans $]1, +\infty[$ ou $]-\infty, -1[$ vérifient

$$x(t) = -\frac{\ln\left|t + \sqrt{t^2 - 1}\right| + C_2}{\sqrt{t^2 - 1}}$$

où C_1 et C_2 sont des constantes. Montrer que cette équation possède une unique solution de classe C^1 sur $]-1, +\infty[$.

EXERCICE 19-1.20 Nous avons vu (exemple **11-3.56**) comment on pouvait parfois obtenir le développement en série entière d'une fonction en utilisant une équation différentielle linéaire. A l'inverse, il est parfois possible de sommer certaines séries entières

par la même méthode :
On considère la série entière

$$\sum_{n=0}^{+\infty} a_n x^{2n+1} = \sum_{n=0}^{+\infty} \frac{n!}{1.3\ldots(2n+1)} x^{2n+1}$$

Montrer que

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{2n+1}$$

et en déduire le rayon de convergence R de cette série entière. L'égalité précédente pouvant s'écrire

$$\forall x \in \mathbb{R} \quad |x| < R \Rightarrow \sum_{n=0}^{+\infty} [(2n+1)\, a_{n+1} - (n+1)\, a_n]\, x^{2n}$$

trouver une équation différentielle linéaire du premier ordre vérifiée sur $]-R, R[$ par la somme de cette série. En déduire que

$$|x| < \sqrt{2} \Rightarrow \sum_{n=0}^{+\infty} \frac{n!}{1.3\ldots(2n+1)} x^{2n+1} = \frac{2}{\sqrt{2-x^2}} \arcsin \frac{x}{\sqrt{2}}$$

REMARQUE 19-1.21 On évite parfois le recours à la méthode de variation des constantes, lorsqu'on trouve une solution "évidente" de l'équation avec second membre. Par exemple, la constante 1 étant solution de

$$(E) : t^2 x' + x = 1$$

les solutions de (E) dans un intervalle ne contenant pas 0 sont de la forme

$$x(t) = 1 + Ce^{\frac{1}{t}}$$

avec C constante arbitraire.

19-1.1.6 Dimension finie : méthode de variation des constantes

Lorsque l'espace $\mathbb{E}_n$ est de dimension $n > 1$, il n'existe pas de "formule" généralisant la proposition **19-1**.14 permettant d'exprimer les solutions de (E_H) à l'aide d'intégrales qui feraient intervenir la fonction $a \in C^0(I, \mathcal{L}(\mathbb{E}_n))$. La démonstration du théorème de Cauchy donne un procédé d'approximation de ces solutions, mais ne donne pas ces dernières explicitement.

Nous nous intéresserons donc ici à la généralisation de la méthode développée à la section précédente, **en supposant connu un système fondamental de solutions de** (E_H)[4]. Si $(x_i)_{1 \leqslant i \leqslant n}$ est une base de l'espace S_H des solutions de (E_H), toute solution de cette équation "sans second membre" s'écrit

$$\sum_{i=1}^{n} C_i\, x_i \quad \text{avec } (C_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n \text{ (constantes)}$$

[4]Pour certaines équations, nous verrons des méthodes particulières d'obtention d'un tel système : notamment l'utilisation de séries entières et, pour les équations "à coefficients constants", la notion d'exponentielle d'une matrice.

Toute fonction $x \in \mathcal{C}^1(I, \mathbb{E}_n)$ peut s'écrire de manière unique sous la forme

$$\forall t \in I \quad x(t) = \sum_{i=1}^{n} C_i(t)\, x_i(t) \quad (*)$$

où $(C_i(t))_{1 \leqslant i \leqslant n}$ est simplement la famille des coordonnées du vecteur $x(t)$ dans la base ("variable") $\mathcal{B}_t = (x_i(t))_{1 \leqslant i \leqslant n}$. On définit ainsi n fonctions $C_i = I \to \mathbb{K}$ qui sont de classe $\mathcal{C}^1$: ceci peut se prouver en utilisant les formules de changement de bases. En effet, si $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ est une base fixe de $\mathbb{E}_n$, la fonction $x \in \mathcal{C}^1(I, \mathbb{E}_n)$ peut se décomposer

$$\forall t \in I \quad x(t) = \sum_{i=1}^{n} \lambda_i(t)\, e_i \quad \text{avec} \quad \lambda_i \in \mathcal{C}^1(I, \mathbb{K})$$

Si $t \in I$, si $X(t) \in \mathcal{GL}_n(\mathbb{K})$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}_t$, ses coefficients sont fonctions de classe $\mathcal{C}^1$ de la variable t. L'égalité

$$\begin{pmatrix} C_1(t) \\ \vdots \\ C_n(t) \end{pmatrix} = [X(t)]^{-1} \begin{pmatrix} \lambda_1(t) \\ \vdots \\ \lambda_n(t) \end{pmatrix}$$

montre qu'il en est de même des fonctions $(C_i)_{1 \leqslant i \leqslant n}$. On peut donc dériver l'équation $(*)$, et on obtient :

$$\forall t \in I \quad x'(t) = \sum_{i=1}^{n} [C_i'(t)\, x_i(t) + C_i(t)\, x_i'(t)]$$

La fonction x est donc solution de (E) si et seulement si

$$\sum_{i=1}^{n} [C_i'(t)\, x_i(t) + C_i(t)\, x_i'(t)] = x'(t)$$

$$= a(t)\, x(t) + b(t) = a(t) \left(\sum_{i=1}^{n} C_i(t)\, x_i(t) \right) + b(t)$$

Comme x_i est solution de (E_H), on a $x_i'(t) = a(t)\, x_i(t)$, ce qui simplifie l'égalité précédente en

$$\sum_{i=1}^{n} C_i'(t)\, x_i(t) = b(t)$$

système d'équations d'inconnues $(C_i'(t))_{1 \leqslant i \leqslant n}$, dont la solution est unique

$$\forall i \quad C_i'(t) = \alpha_i(t) \quad i^{\text{ème}} \text{ coordonnée de } b(t) \text{ dans } \mathcal{B}_t$$

Si A_i est une primitive de la fonction (continue) α_i sur I, on aura

$$\forall t \in I \quad C_i(t) = A_i(t) + \lambda_i \quad \text{avec } \lambda_i \in \mathbb{K}$$

et donc

$$\forall t \in I \quad x(t) = \sum_{i=1}^{n} A_i(t)\, x_i(t) + \sum_{i=1}^{n} \lambda_i x_i(t)$$

ce qui est conforme à la structure d'espace affine de l'ensemble des solutions de (E).

EXEMPLE 19-1.22 Résoudre[5] le système d'équations scalaires réelles

$$(E) : \begin{cases} x_1'(t) = (t-1)\, x_1(t) + (2-t)\, e^t x_2(t) + \dfrac{1}{t} \\ x_2'(t) = (t-1)\, e^{-t} x_1(t) + (1-t)\, x_2(t) + \dfrac{e^{-t}}{1-t} \end{cases}$$

dans un intervalle I ne contenant ni 0 ni 1.

On travaille donc ici dans $\mathbb{E}_2 = \mathbb{R}^2$ rapporté à sa base canonique, avec le vecteur "inconnu"

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

On remarque que, en posant $z(t) = e^t x_2(t)$, le système homogène associé à (E_H) équivaut à

$$(E_H) \Leftrightarrow \begin{cases} x_1'(t) = (t-1)\, x_1(t) + (2-t)\, z(t) \\ z'(t) = (t-1)\, x_1(t) + (2-t)\, z(t) \end{cases}$$

ce qui montre que la différence $x_1 - z$ est constante, et si λ est sa valeur, la première équation donne

$$x_1'(t) = x_1(t) + \lambda(t-2)$$

Les solutions de cette équation s'obtiennent comme on l'a vu à la section précédente, ou en remarquant simplement que $t \mapsto \lambda(1-t)$ est solution particulière de l'équation avec second membre. On a donc

$$x_1(t) = \lambda(1-t) + \mu e^t$$

ce qui donne $x_2(t) = e^{-t} z(t) = e^{-t}(x_1(t) - \lambda) = -\lambda t\, e^{-t} + \mu$, soit

$$X(t) = -\lambda \begin{pmatrix} t-1 \\ te^{-t} \end{pmatrix} + \mu \begin{pmatrix} e^t \\ 1 \end{pmatrix}$$

On obtient le système fondamental de solutions

$$X_1(t) = \begin{pmatrix} t-1 \\ te^{-t} \end{pmatrix} \quad \text{et} \quad X_2(t) = \begin{pmatrix} e^t \\ 1 \end{pmatrix}$$

[5]Il s'agit évidemment d'un exemple "construit sur mesure", pour lequel on peut trouver un système fondamental de solutions de l'équation sans second membre.

On vérifie la formule de Liouville sur cet exemple : le wronskien de ces deux solutions dans la base canonique vaut

$$\begin{vmatrix} t-1 & e^t \\ te^{-t} & 1 \end{vmatrix} = -1$$

Il est constant puisque la matrice du système

$$A(t) = \begin{pmatrix} t-1 & (2-t)\,e^t \\ (t-1)\,e^{-t} & 1-t \end{pmatrix}$$

est de trace nulle.
Les solutions de (E) pourront donc être cherchées sous la forme

$$X(t) = C_1(t)\,X_1(t) + C_2(t)\,X_2(t) \quad \text{avec } C_1 \text{ et } C_2 \in C^1(I,\mathbb{R})$$

On obtient

$$C_1'(t)\,X_1(t) + C_2'(t)\,X_2(t) = B(t) = \begin{pmatrix} \dfrac{1}{t} \\ \dfrac{e^{-t}}{1-t} \end{pmatrix}$$

système d'inconnues $C_1'(t)$ et $C_2'(t)$ qu'on peut résoudre par exemple en utilisant les déterminants

$$C_1'(t) = -\begin{vmatrix} \dfrac{1}{t} & e^t \\ \dfrac{e^{-t}}{1-t} & 1 \end{vmatrix} = -\left(\frac{1}{t} + \frac{1}{t-1}\right) \text{ et } C_2'(t) = -\begin{vmatrix} t-1 & \dfrac{1}{t} \\ te^{-t} & \dfrac{e^{-t}}{1-t} \end{vmatrix} = 2e^{-t}$$

ce qui permet ensuite d'obtenir $C_1(t)$ et $C_2(t)$

$$C_1(t) = -\ln|t(t-1)| + \alpha \quad \text{et} \quad C_2(t) = -2e^{-t} + \beta$$

où α et β sont deux constantes.On obtient finalement

$$x_1(t) = (1-t)\ln|t(t-1)| - 2 + \alpha(t-1) + \beta e^t$$

$$x_2(t) = -te^{-t}\ln|t(t-1)| - 2e^{-t} + \alpha te^{-t} + \beta$$

EXERCICE 19-1.23 Résoudre le système différentiel

$$\begin{cases} x_1'(t) = tx_1(t) - x_2(t) + e^{\frac{t^2}{2}} \\ x_2'(t) = x_1(t) + tx_2(t) + te^{\frac{t^2}{2}} \end{cases}$$

Pour résoudre l'équation homogène, on pourra remarquer que la matrice du système

$$A(t) = \begin{pmatrix} t & -1 \\ 1 & t \end{pmatrix}$$

est une matrice de similitude directe, donc d'une transformation géométrique qui s'exprime simplement lorsqu'on passe en affixe complexe. On pourra donc poser

$$z(t) = x_1(t) + ix_2(t)$$

19-1.2 Cas des équations à coefficients constants

On étudie ici le cas particulier où l'application $t \mapsto a(t)$ est constante : l'endomorphisme $a \in \mathcal{L}_c(\mathbb{E})$ est indépendant de t et, dans le cas de la dimension finie, sa matrice $A \in \mathcal{M}_n(\mathbb{K})$ représentative dans une base de $\mathbb{E}_n$ a tous ses coefficients constants.

19-1.2.1 Résolution de l'équation homogène

Rappelons le résultat obtenu à la section **9-6.3.3**

PROPOSITION 19-1.24 Si $a \in \mathcal{L}_c(\mathbb{E})$, les solutions (définies sur $\mathbb{R}$) de l'équation différentielle homogène

$$(E_H) \quad x' = ax$$

sont de la forme

$$x(t) = \exp(ta)\, z$$

où $z \in \mathbb{E}$ est un vecteur arbitraire. En particulier, l'unique solution de (E_H) vérifiant la condition initiale $x(t_0) = x_0$ est donnée par

$$x(t) = \exp((t - t_0)\, a)\, x_0$$

Rappelons que ce résultat a été obtenu par "variation de la constante", en montrant simplement qu'une fonction $x \in C^1(\mathbb{R}, \mathbb{E})$ est solution de (E_H) si et seulement si la fonction $t \mapsto y(t) = \exp(-ta)\, x(t)$ est constante. La formule précédente montre aussi que, dans le cas particulier où $\mathbf{E}_n = \mathbb{K}^n$ avec a identifié à sa matrice A dans la base canonique, les vecteurs colonnes de la matrice $\exp(tA)$ définissent n fonctions de $\mathbb{R}$ dans $\mathbb{K}^n$ qui forment un système fondamental de solutions de (E_H).

REMARQUE 19-1.25 Dans ce cas, pour t_0 et $t_1 \in \mathbb{R}$, les automorphismes $(R(t_0, t_1))_{(t_0,t_1)\in\mathbb{R}^2}$ de $\mathbb{E}$ définis à la remarque **19-1.**8 sont simplement donnés par

$$R(t_0, t_1) = \exp((t_1 - t_0)\, a)$$

et la propriété de composition de ces automorphismes s'écrit

$$\forall s, t \in \mathbb{R} \quad \exp((t + s)\, a) = \exp(ta) \exp(sa)$$

De même, en dimension finie, la formule de Liouville (*cf.* proposition **19-1.**12) peut alors se formuler, pour une base $\mathcal{B}_0 = (e_i)_{1 \leqslant i \leqslant n}$ quelconque de $\mathbf{E}_n$

$$\forall t \in \mathbb{R} \quad \det_{\mathcal{B}_0}(\exp(ta)\, e_1, \ldots, \exp(ta)\, e_n) = \det(\exp(ta)) = e^{t \operatorname{trace} a}$$

COROLLAIRE 19-1.26 Si x_0 est un vecteur propre de a pour une valeur propre λ, la solution de (E_H) vérifiant la condition initiale $x(t_0) = x_0$ est donnée par

$$x(t) = e^{\lambda(t-t_0)} x_0$$

Démonstration : On peut utiliser la formule de la proposition précédente ou, plus simplement, voir que la fonction x définie par $x(t) = e^{\lambda(t-t_0)}x_0$ vérifie $x(t_0) = x_0$ et

$$x'(t) = \lambda e^{\lambda(t-t_0)}x_0 = a\left(e^{\lambda(t-t_0)}x_0\right)$$

puisque x_0 est vecteur propre de a pour la valeur propre λ. ∎

Nous supposerons dans la suite de cette section que l'espace $\mathbb{E} = \mathbb{E}_n$ est de dimension finie[6].

COROLLAIRE 19-1.27 Si $\mathbb{F}$ est un sous-espace vectoriel de $\mathbb{E}_n$ stable par a, pour toute solution x de (E_H) on a

$$\exists t_0 \in \mathbb{R} \quad x(t_0) \in \mathbb{F} \Leftrightarrow \forall t \in \mathbb{R} \quad x(t) \in \mathbb{F}$$

Démonstration : Si $x(t_0) = x_0 \in \mathbb{F}$, on a

$$\forall t \in \mathbb{R} \quad x(t) = \exp((t-t_0)a)\, x_0$$

et $\mathbb{F}$ étant stable par l'endomorphisme $u = (t-t_0)a$ l'est aussi par $\exp(u)$. On peut aussi éviter le recours à l'expression exacte de la solution, en utilisant le théorème de Cauchy : si $b \in \mathcal{L}(\mathbb{F})$ est l'endomorphisme induit par a sur $\mathbb{F}$, l'équation

$$y' = by$$

possède une unique solution $y \in \mathcal{C}^1(\mathbb{R}, \mathbb{F})$ vérifiant $y(t_0) = x_0$. C'est clairement une solution de (E_H) vérifiant la même condition initiale que x, et elle est donc égale à x. ∎

En d'autres termes, une solution de (E_H) qui prend une valeur dans l'espace $\mathbb{F}$ a sa trajectoire totalement incluse dans $\mathbb{F}$. La donnée d'une condition initiale correspondant à un vecteur propre de a est évidemment un cas particulier de ce résultat (avec $\mathbb{F}$ droite vectorielle).

COROLLAIRE 19-1.28 Si $a \in \mathcal{L}(\mathbb{E}_n)$ est diagonalisable, et si $(u_i)_{1\leqslant i\leqslant n}$ est une base de vecteurs propres pour les valeurs propres $(\lambda_i)_{1\leqslant i\leqslant n}$ (non nécessairement distinctes deux à deux), la famille de fonctions $(x_i)_{1\leqslant i\leqslant n}$ définies par

$$x_i(t) = e^{\lambda_i t}u_i$$

forme un système fondamental de solutions de l'équation

$$(E_H) \quad x' = ax$$

[6]Certains résultats subsistent en dimension infinie, notamment le corollaire **19-1**.27 si on suppose le sous-espace $\mathbb{F}$ fermé.

Démonstration : C'est une famille de n solutions telle que la famille $(x_i(0))_{1\leqslant i\leqslant n} = (u_i)_{1\leqslant i\leqslant n}$ est une base de $\mathbb{E}_n$. Si un vecteur $x_0 \in \mathbb{E}_n$ se décompose dans la base $(u_i)_{1\leqslant i\leqslant n}$ sous la forme

$$x_0 = \sum_{i=1}^{n} \alpha_i u_i$$

la solution x vérifiant la condition initiale $x(0) = x_0$ sera alors donnée par

$$\forall t \in \mathbb{R} \quad x(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} u_i \quad \blacksquare$$

19-1.2.2 Etude en dimension 2

Nous allons illustrer les résultats précédents dans le cas d'un espace vectoriel réel $\mathbb{E}_2$ de dimension 2. On considère un endomorphisme a de $\mathbb{E}_2$ et l'équation $(E_H) : \vec{x}' = a\vec{x}$. Si on travaille dans une base $\mathcal{B}_0 = (\vec{e}_1, \vec{e}_2)$ quelconque de $\mathbb{E}_2$, où a est représenté par une matrice $A \in \mathcal{M}_2(\mathbb{R})$, (E_H) se traduira par le système

$$(S_H) : X'(t) = AX(t) \Leftrightarrow \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

Nous supposons $a \neq 0_{\mathcal{L}(\mathbb{E}_2)}$. La discussion va se mener en étudiant d'abord le cas où a est diagonalisable (c'est à dire A est $\mathbb{R}$-diagonalisable), puis le cas où A possède deux valeurs propres complexes conjuguées (distinctes !) et enfin le cas où a est trigonalisable (une valeur propre réelle double, a n'étant pas une homothétie).

1. **Cas où a est diagonalisable.**

   Si $(\vec{u}_1, \vec{u}_2)$ est une base de vecteur propres de a pour les valeurs propres (λ_1, λ_2), le cas $\lambda_1 = \lambda_2 = \lambda \neq 0$ correspond au cas très particulier où a est une homothétie. Les solutions de (E_H) sont alors toutes de la forme

   $$\vec{x}(t) = e^{\lambda t}\vec{x}_0$$

   et les trajectoires sont alors des demi-droites issues de l'origine (figure **19.2**), (sauf pour la solution identiquement nulle dont la trajectoire est réduite à un point, correspondant à la condition initiale $\vec{x}_0 = \vec{0}_{\mathbb{E}_2}$). On dit que l'origine est un point d'équilibre[7] pour l'équation $\vec{x}' = a\vec{x}$. Pour $t \to +\infty$, toutes les solutions tendent vers $\vec{0}_{\mathbb{E}_2}$ si $\lambda < 0$, l'équilibre est alors dit (asymptotiquement) stable. Lorsque $\lambda > 0$ l'équilibre est instable, toutes les solutions (non nulles) vérifiant

   $$\lim_{t\to+\infty} \|\vec{x}(t)\| = +\infty$$

[7] Situation générale pour toutes les équations différentielles linéaires homogènes.

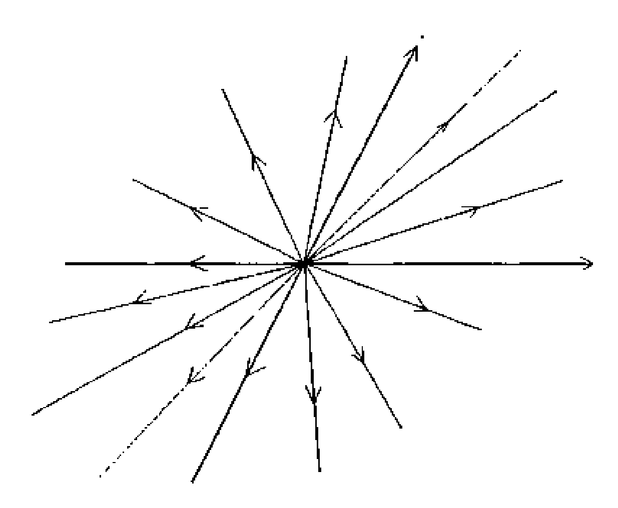


Figure **19.2** – *Cas où a est une homothétie* $\lambda_1 = \lambda_2 > 0$

Nous considérons donc à présent le cas $\lambda_1 \neq \lambda_2$, et nous pouvons supposer qu'au moins une des valeurs propres de a est strictement positive : en effet, si $t \mapsto \vec{x}(t)$ est une solution de $\vec{x}' = a\vec{x}$, la fonction $t \mapsto \vec{x}(-t)$ est solution de $\vec{x}' = (-a)\vec{x}$. Les trajectoires sont les mêmes, seul le sens de parcours est inversé. On supposera donc $\lambda_1 > \lambda_2$ avec $\lambda_1 > 0$. Les solutions de (E_H) sont toutes de la forme

$$\vec{x}(t) = \alpha_1 e^{\lambda_1 t}\vec{u}_1 + \alpha_2 e^{\lambda_2 t}\vec{u}_2$$

Dans le plan rapporté à un repère $(O, \vec{u}_1, \vec{u}_2)$, le point $M(t)$ vérifiant l'égalité $\overrightarrow{OM}(t) = \vec{x}(t)$ a donc pour coordonnées

$$M(t) : \begin{pmatrix} x_1(t) = \alpha_1 e^{\lambda_1 t} \\ x_2(t) = \alpha_2 e^{\lambda_2 t} \end{pmatrix}$$

Les axes de coordonnées se décomposent en général en deux demi-droites ouvertes et en l'origine, qui correspondent à des trajectoires particulières[8]. L'allure des autres trajectoires (pour lesquelles α_1 et $\alpha_2 \neq 0$) dépend alors du signe de λ_2 :

- Si $\lambda_2 > 0$, on a

  $$\lim_{t \to +\infty} x_1(t) = \pm\infty \text{ et } \lim_{t \to +\infty} x_2(t) = \pm\infty \quad \text{avec} \quad x_2(t) \underset{t \to +\infty}{=} o(x_1(t))$$

  Les trajectoires présentent des branches infinies avec direction asymptotique égale à $\mathrm{vect}(\vec{u}_1)$. Par contre

  $$\lim_{t \to -\infty} x_1(t) = \lim_{t \to -\infty} x_2(t) = 0 \quad \text{avec} \quad x_1(t) \underset{t \to +\infty}{=} o(x_2(t))$$

[8]Ces trajectoires correspondent aux solutions

$$t \mapsto (\pm e^{\lambda_1 t}, 0) \quad \text{et} \quad t \mapsto (0, \pm e^{\lambda_2 t})$$

Dans le cas particulier où $\lambda_2 = 0$, tous les points de l'axe des ordonnées sont points d'équilibre (solutions correspondant à une condition initiale dans le noyau de l'endomorphisme a).

Donc, pour $t \to -\infty$, le point $M(t)$ tend vers l'origine, et la trajectoire ainsi prolongée possède une tangente dirigée par le vecteur $\vec{u}_2$ (figure **19.3**).

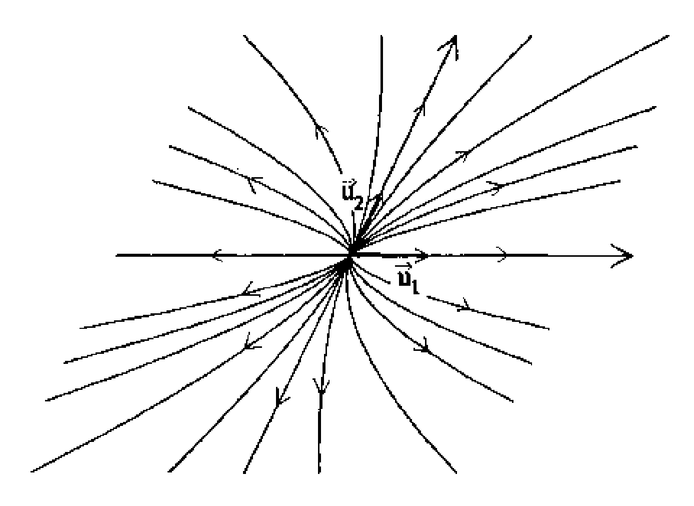


Figure **19.3** – *Cas où a est diagonalisable* $\lambda_1 > \lambda_2 > 0$

- Si $\lambda_2 = 0$, les trajectoires sont des demi-droites parallèles au vecteur $\vec{u}_1$ (figure **19.4**).

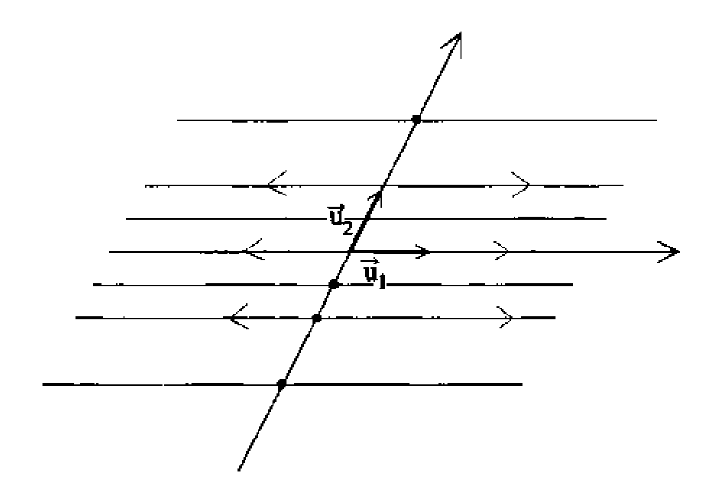


Figure **19.4** – *Cas où a est diagonalisable* $\lambda_1 > \lambda_2 = 0$

- Si $\lambda_2 < 0$, on a

$$\lim_{t\to+\infty} x_1(t) = \pm\infty \quad \text{et} \quad \lim_{t\to+\infty} x_2(t) = 0$$
$$\lim_{t\to-\infty} x_1(t) = 0 \quad \text{et} \quad \lim_{t\to+\infty} x_2(t) = \pm\infty$$

et les trajectoires présentent deux asymptotes correspondant aux axes de coordonnées (figure **19.5**).

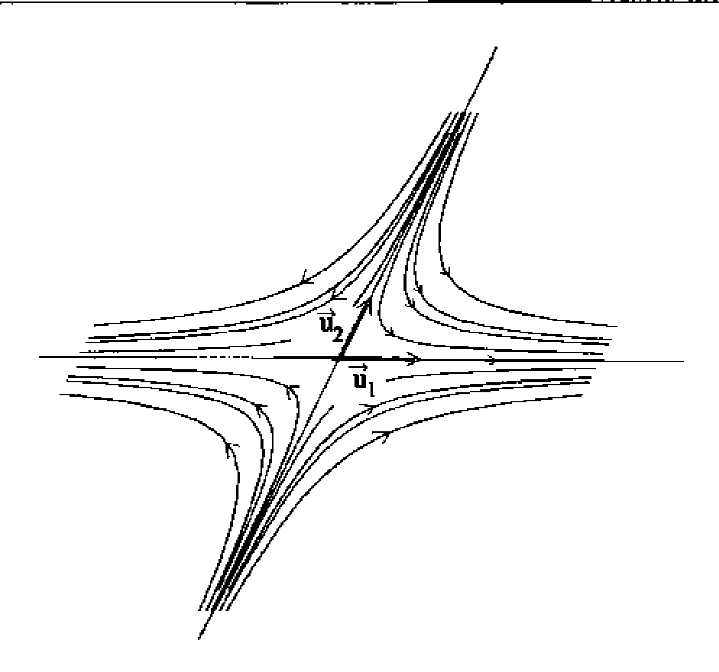


Figure **19.5** – *Cas où a est diagonalisable* $\lambda_1 > 0 > \lambda_2$

2. **Cas où la matrice** A **possède deux valeurs propres conjuguées.**

 On a $\lambda_1 = \alpha + i\beta$ et $\lambda_2 = \alpha - i\beta$ avec $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^*$. Si $X_1 \in \mathbb{C}^2$ est vecteur propre de A pour λ_1, son conjugué $\overline{X_1}$ est vecteur propre pour λ_2. Les fonctions

$$\varphi_1 : t \mapsto e^{(\alpha+i\beta)t} X_1 \quad \text{et} \quad \varphi_2 : t \mapsto e^{(\alpha-i\beta)t} \overline{X_1}$$

 forment donc un système fondamental de solutions complexes du système différentiel $X' = AX$. Le système qui s'en déduit par combinaisons linéaires

$$\psi_1 = \frac{\varphi_1 + \varphi_2}{2} : t \mapsto \mathrm{Re}\left(e^{(\alpha+i\beta)t} X_1\right) \text{ et } \psi_1 = \frac{\varphi_1 - \varphi_2}{2i} : t \mapsto \mathrm{Im}\left(e^{(\alpha+i\beta)t} X_1\right)$$

 est donc clairement un système fondamental de solutions réelles du même système. En séparant X_1 en parties réelle et imaginaire

$$X_1 = Y_1 + iY_2$$

 on aura donc le système fondamental

$$t \mapsto e^{\alpha t} (\cos \beta t \, Y_1 - \sin \beta t \, Y_2) \quad \text{et} \quad t \mapsto e^{\alpha t} (\sin \beta t \, Y_1 + \cos \beta t \, Y_2)$$

 Si $\mathcal{B}_0 = (\vec{e}_1, \vec{e}_2)$ est la base de $\mathbb{E}_2$ où a est représenté par la matrice A, et si $\vec{v}_1$ et $\vec{v}_2$ sont les vecteurs (indépendants[9]) représentés dans cette base par les colonnes Y_1 et Y_2, les solutions de $\vec{x}' = a\vec{x}$ s'écriront alors sous la forme

$$\vec{x}(t) = e^{\alpha t} (\lambda \cos \beta t + \mu \sin \beta t) \, \vec{v}_1 + e^{\alpha t} (-\lambda \sin \beta t + \mu \cos \beta t) \, \vec{v}_2$$

[9] On vérifie facilement que si, dans $\mathbb{C}^n$, un vecteur et son conjugué sont $\mathbb{C}$-indépendants, la partie réelle et la partie imaginaire de l'un de ces vecteurs sont indépendantes dans $\mathbb{R}^n$.

où λ et μ sont deux constantes réelles, ou encore, par transformation d'expressions trigonométriques

$$\vec{x}(t) = Me^{\alpha t}\left[\cos\left(\beta\left(t-\varphi\right)\right)\vec{v}_1 - \sin\left(\beta\left(t-\varphi\right)\right)\vec{v}_2\right]$$

avec M et φ constantes. L'allure des trajectoires dépend alors de α :

- Si $\alpha < 0$, on obtient des spirales s'enroulant autour de l'origine, qui est un point d'équilibre asymptotiquement stable.
- Si $\alpha > 0$, on obtient un résultat analogue, avec les trajectoires parcourues en sens inverse, l'origine est à présent instable (figure **19.6**).

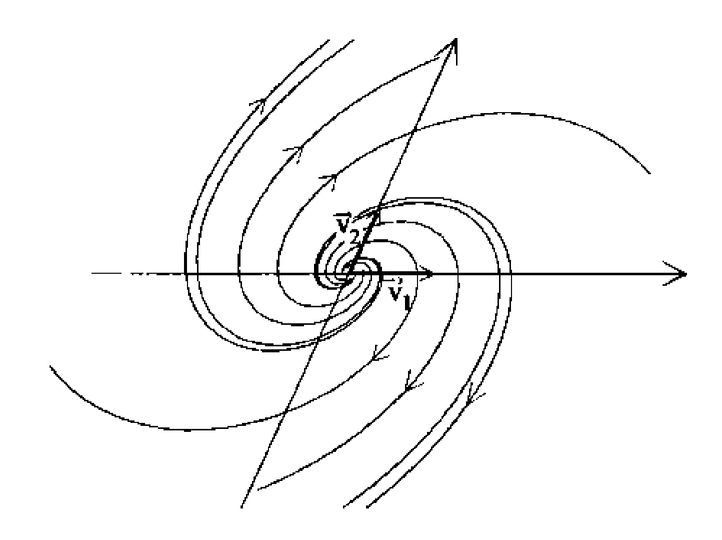


Figure **19.6** – *Cas où a est $\mathbb{C}$-diagonalisable $\lambda = \alpha \pm i\beta$ avec $\alpha > 0$*

- Si $\alpha = 0$, les trajectoires sont des ellipses centrées en O. L'origine est alors un point d'équilibre stable en ce sens qu'une condition initiale "petite" va donner une trajectoire complètement incluse dans un voisinage de O fixé arbitrairement. Il ne s'agit plus ici de stabilité asymptotique (figure **19.7**).

3. **Cas où a possède une valeur propre double.**

   On suppose ici que a n'est pas une homothétie. Il existe alors une base $(\vec{u}_1, \vec{u}_2)$ de $\mathbf{E}_2$ où la matrice de a s'écrit

$$A_1 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

   et si $X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ est la colonne des coordonnées de $\vec{x}(t)$ dans cette base, le système (E_H) est équivalent à

$$\begin{cases} x_1'(t) = \lambda x_1(t) + x_2(t) \\ x_2'(t) = \lambda x_2(t) \end{cases}$$

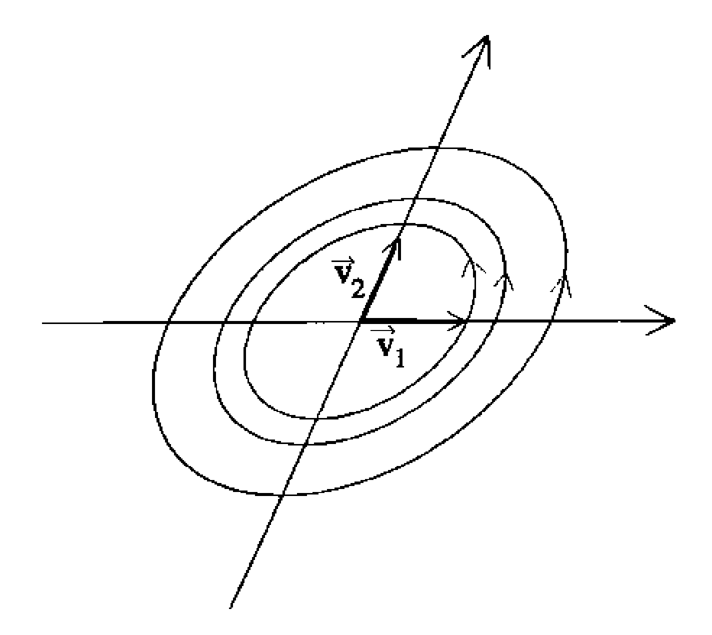


Figure **19.7** – *Cas où a est* $\mathbb{C}$*-diagonalisable* $\lambda = \pm i\beta$

La seconde équation donne alors $x_2(t) = Ce^{\lambda t}$ avec $C \in \mathbb{R}$ constante. En reportant dans la première équation, on obtient

$$x_1'(t) = \lambda x_1(t) + Ce^{\lambda t}$$

dont les solutions s'obtiennent en cherchant une solution particulière de la forme $t \mapsto \mu t e^{\lambda t}$, ce qui donne

$$x_1'(t) = e^{\lambda t}(\mu + \lambda\mu t) = \lambda x_1(t) + Ce^{\lambda t} = \lambda\mu t e^{\lambda t} + Ce^{\lambda t}$$

ce qui donne $\mu = C$. On obtient finalement la solution générale en ajoutant à cette solution particulière une fonction $t \mapsto Be^{\lambda t}$ avec B constante. On arrive à

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = B\begin{pmatrix} e^{\lambda t} \\ 0 \end{pmatrix} + C\begin{pmatrix} te^{\lambda t} \\ e^{\lambda t} \end{pmatrix}$$

ce qui permet d'écrire les solutions sous la forme

$$\vec{x}(t) = (B + Ct)e^{\lambda t}\,\vec{u}_1 + Ce^{\lambda t}\,\vec{u}_2$$

Pour $C = 0$, on retrouve évidemment les trajectoires correspondant à une condition initiale colinéaire à $\vec{u}_1$, vecteur propre de a.

L'allure des courbes intégrales dépend de la position de λ par rapport à 0 :

- Si $\lambda > 0$ (le cas $\lambda < 0$ s'étudie de façon analogue en renversant le temps), les trajectoires pour lesquelles $C \neq 0$ vérifient

$$\lim_{t\to+\infty} x_1(t) = \pm\infty \text{ et } \lim_{t\to+\infty} x_2(t) = \pm\infty \quad \text{avec} \quad \lim_{t\to+\infty}\left|\frac{x_1(t)}{x_2(t)}\right| = +\infty$$

ce qui correspond à une branche infinie avec direction asymptotique égale à vect $(\vec{u}_1)$. On a également

$$\lim_{t\to-\infty} x_1(t) = 0 \text{ et } \lim_{t\to-\infty} x_2(t) = 0 \quad \text{avec} \quad \lim_{t\to+\infty} \left|\frac{x_1(t)}{x_2(t)}\right| = 0$$

Les courbes intégrales se prolongent donc jusqu'à l'origine, avec une tangente dirigée par $\vec{u}_1$ (figure **19.8**).

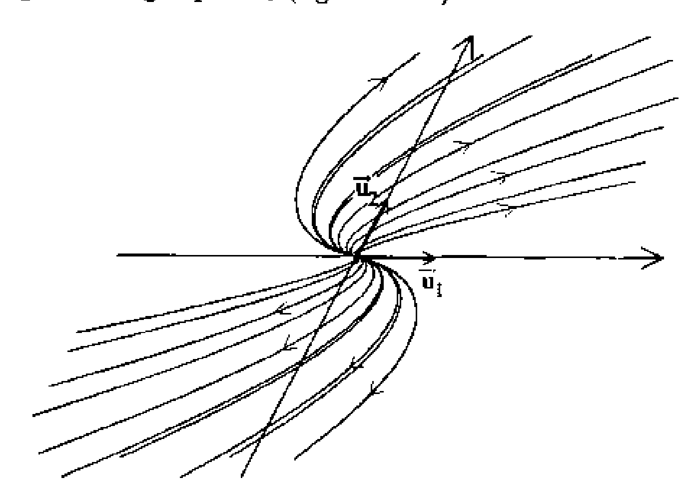


Figure **19.8** – *Cas où a possède une valeur propre double $\lambda > 0$*

- Si $\lambda = 0$, les trajectoires sont des droites dirigées par le vecteur $\vec{u}_1$ (figure **19.9**).

19-1.2.3 Résolution pratique

On suppose dans cette section que $\mathbb{E}_n = \mathbb{K}^n$, situation à laquelle on peut toujours se ramener en travaillant dans une base de $\mathbb{E}_n$. L'équation (E_H) s'écrit alors

$$X' = AX$$

avec $A \in \mathcal{M}_n(\mathbb{K})$. L'expression théorique des solutions

$$X(t) = \exp(tA)\, X_0$$

montre qu'en général, les calculs explicites des solutions seront pénibles. Deux cas plus simples sont à retenir :

- **Si A est diagonalisable** et si on connaît une base de vecteurs propres $(X_i)_{1\leqslant i\leqslant n}$ pour les valeurs propres $(\lambda_i)_{1\leqslant i\leqslant n}$, les solutions de (E_H) seront de la forme

$$X(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} X_i$$

où les $(\alpha_i)_{1\leqslant i\leqslant n}$ sont des constante. Remarquons que, lorsque $\mathbb{K} = \mathbb{R}$ et que la matrice A est $\mathbb{C}$-diagonalisable, on peut obtenir un système fondamental de solutions réelles de (E_H) comme nous l'avons fait dans le cas $n = 2$, en séparant les parties réelles et imaginaires des solutions complexes.

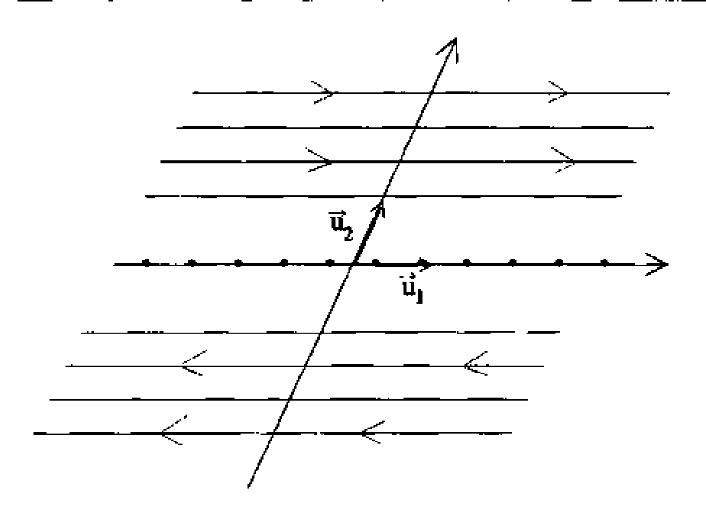


Figure **19.9** – *Cas où a possède une valeur propre double* $\lambda = 0$

- **Si A n'a qu'une seule valeur propre** λ, on sait qu'on peut alors écrire

$$\forall t \in \mathbb{R} \quad tA = \lambda t I_n + tN$$

avec N nilpotente, ce qui donnera

$$\exp(tA) = e^{\lambda t} \exp(tN) = e^{\lambda t} \left(\sum_{k=0}^{n-1} \frac{t^k}{k!} N^k \right)$$

Les solutions sont de la forme

$$X(t) = e^{\lambda t} \left(\sum_{k=0}^{n-1} \frac{t^k}{k!} N^k X_0 \right)$$

vecteur colonne dont les coordonnées sont des fonctions de la forme $t \mapsto e^{\lambda t} Q(t)$, avec $Q \in \mathbb{K}_{n-1}[X]$.

- **Dans le cas général**, on peut bien sûr calculer $\exp(tA)$ (en utilisant la réduction dans les sous-espaces caractéristiques, voir la section **9-6.3.2**). On peut aussi utiliser la méthode des "coefficients indéterminés", basée sur les remarques suivantes :

  On considère le système dans $\mathbb{C}^n$, puisque d'un système fondamental de solutions complexes on peut déduire, si $A \in \mathcal{M}_n(\mathbb{R})$, un système fondamental réel. L'espace $\mathbb{C}^n$ est donc somme directe des sous-espaces caractéristiques de A

$$\mathbb{C}^n = \bigoplus_{i=1}^{p} \ker (A - \lambda_i I_n)^{n_i}$$

  si $(\lambda_i)_{1 \leqslant i \leqslant p}$ sont les valeurs propres de A, de multiplicités respectives $(n_i)_{1 \leqslant i \leqslant p}$. Si une condition initiale $X_0 \in \mathbb{C}^n$ se décompose en

$$X_0 = \sum_{i=1}^{p} Y_i$$

dans cette somme directe, on aura pour la solution correspondante

$$X(t) = \exp(tA)\,X_0 = \sum_{i=1}^{p} \exp(tA)\,Y_i \quad (*)$$

Lorsque Y_i parcourt le sous-espace caractéristique associé à λ_i, les solutions $t \to \exp(tA)\,Y_i$ décrivent un espace vectoriel de dimension n_i, et ces solutions sont de la forme

$$t \mapsto e^{\lambda_i t}(Z_0 + tZ_1 + \cdots + t^{n_i-1}Z_{n_i-1})$$

(les Z_j sont dans le sous-espace caractéristique associé à la valeur propre λ_i), d'après l'étude effectuée dans la cas d'un endomorphisme n'ayant qu'une seule valeur propre et la proposition **19-1**.27. Il existe donc un sous-espace $\mathcal{S}_H^i$ de $\mathcal{S}_H$, de dimension n_i, composé de fonctions de la forme précédente. On peut trouver ces solutions en résolvant un système linéaire dont les inconnues sont les composantes des vecteurs $(Z_j)_{1\leqslant j\leqslant n_i-1}$. La somme directe des $(\mathcal{S}_H^i)_{1\leqslant i\leqslant p}$ est égale à $\mathcal{S}_H$, d'après $(*)$.

EXEMPLE 19-1.29 Trouver un système fondamental de solutions de

$$X' = \begin{pmatrix} 0 & 2 & 2 \\ -1 & 2 & 2 \\ -1 & 1 & 3 \end{pmatrix} X$$

On vérifie le polynôme caractéristique de la matrice A de ce système vaut

$$\chi_A(X) = (1-X)(2-X)^2$$

Un vecteur propre pour la valeur propre 1 est le vecteur $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, ce qui donne déjà la solution

$$t \mapsto e^t \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

On cherche ensuite des solutions de la forme

$$X(t) = e^{2t} \begin{pmatrix} a_1t + b_1 \\ a_2t + b_2 \\ a_3t + b_3 \end{pmatrix}$$

et, par identification de coefficients, on tombe sur les systèmes

$$\begin{cases} a_1 = a_2 + a_3 \\ 2a_2 = -a_1 + 2a_2 + 2a_3 \\ 2a_3 = -a_1 + a_2 + 3a_3 \end{cases} \quad \text{et} \quad \begin{cases} a_1 + 2b_1 = 2b_2 + 2b_3 \\ a_2 + 2b_2 = -b_1 + 2b_2 + 2b_3 \\ a_3 + 2b_3 = -b_1 + b_2 + 3b_3 \end{cases}$$

Le premier système donne

$$\exists \mu \in \mathbb{R} \quad (a_1, a_2, a_3) = \mu(2,1,1)$$

et en reportant dans le second, on obtient

$$\exists \nu \in \mathbb{R} \quad (b_1, b_2, b_3) = (-\mu + 2\nu, \nu, \nu)$$

ce qui donne finalement le système fondamental de solutions $(X_i)_{1 \leqslant i \leqslant 3}$ avec

$$X_1(t) = e^t \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},\ X_2(t) = e^{2t} \begin{pmatrix} 2t-1 \\ t \\ t \end{pmatrix} \text{ et } X_3(t) = e^{2t} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Ce calcul montre en particulier que le vecteur de composantes $(2,1,1)$ est vecteur propre de A pour la valeur propre 2, et que le vecteur $(1,0,0)$ est aussi dans le sous-espace caractéristique correspondant.

REMARQUE 19-1.30 Si on dispose d'une matrice de passage P trigonalisant la matrice A, c'est à dire telle que

$$A = PTP^{-1}$$

avec T triangulaire supérieure, on peut aussi résoudre le système $X' = AX$ par changement de fonction inconnue : si $Y = P^{-1}X$, on a

$$Y' = P^{-1}X' = TP^{-1}X = TY$$

système qu'on peut résoudre en intégrant n équations différentielles scalaires : si

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

la dernière équation du système $Y' = TY$ détermine y_n (c'est une fonction exponentielle). En reportant dans l'équation qui précède, on obtient y_{n-1} en résolvant une équation linéaire avec second membre, et ainsi de suite. On revient ensuite à X par $X = PY$ (et on n'a pas besoin de calculer P^{-1}).

EXERCICE 19-1.31 Montrer que toutes les solutions du système $X' = AX$ vérifient

$$\lim_{t \to +\infty} X(t) = 0$$

si et seulement si toutes les valeurs propres de A ont une partie réelle strictement négative. A quelle condition les solutions sont-elles toutes bornées sur $]0, +\infty[$?

19-1.2.4 Equation avec second membre

Dans le cas d'un système à coefficients constants, la méthode de variation des constantes donne une expression intégrale simple (en théorie) des solutions de l'équation avec second membre

$$(E) : x' = ax + b$$

où $b \in \mathcal{C}^0(I, \mathbb{E}_n)$. On cherche la solution vérifiant la condition initiale

$$x(t_0) = x_0$$

On fait alors changement de fonction inconnue

$$y(t) = e^{-ta} x(t)$$

ce qui revient à "poser" $x(t) = e^{ta} y(t)$. En reportant cette expression dans (E), on obtient

$$y'(t) = e^{-ta} b(t)$$

ce qui donne

$$\forall t \in I \quad y(t) = y(t_0) + \int_{t_0}^{t} e^{-ua} b(u)\, du$$

et donc

$$\forall t \in I \quad x(t) = e^{ta} \left(e^{-t_0} x_0 + \int_{t_0}^{t} e^{-ua} b(u)\, du \right)$$

soit également

$$\forall t \in I \quad x(t) = e^{(t-t_0)a} x_0 + \int_{t_0}^{t} e^{(t-u)a} b(u)\, du$$

Cette formule est essentiellement théorique. Si on dispose d'une solution particulière de l'équation avec second membre, on utilisera plutôt la structure d'espace affine de l'ensemble des solutions. Ce sera notamment le cas pour un système écrit matriciellement

$$(E) : X'(t) = AX(t) + B(t)$$

lorsque le second membre est de la forme

$$B(t) = e^{\mu t} P(t)$$

où $\mu \in \mathbb{K}$ et $P(t)$ est un vecteur colonne dont les composantes sont des fonctions polynômes. On démontre en effet que, si tous les polynômes composant P sont de degré $\leqslant p$, (E) possède une solution particulière de la forme

$$X(t) = e^{\mu t} Q(t)$$

où les composantes de $Q(t)$ sont des fonctions polynômes de degré $\leqslant p + m$, avec m égal à la multiplicité de μ comme valeur propre de A ($m = 0$ si μ n'est pas valeur propre de A).

EXEMPLE 19-1.32 Déterminer les solutions réelles du système

$$\begin{cases} x'(t) = -2x(t) + 6y(t) - 2z(t) \\ y'(t) = -x(t) + 3y(t) - z(t) - 2t \\ z'(t) = -x(t) + 3y(t) + z(t) - 5\sin t \end{cases}$$

Les valeurs propres de la matrice

$$A = \begin{pmatrix} -2 & 6 & -2 \\ -1 & 3 & -1 \\ -1 & 3 & 1 \end{pmatrix}$$

sont 0 et $1 \pm i$, associées aux vecteurs propres $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} \pm 2i \\ \pm i \\ 1 \end{pmatrix}$. On obtient donc un système fondamental de solutions : si $Z(t) = e^{(1+i)t}\begin{pmatrix} 2i \\ i \\ 1 \end{pmatrix}$, on peut prendre

$$X_1(t) = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix},\ X_2(t) = \operatorname{Re} Z(t) = e^t \begin{pmatrix} -2\sin t \\ -\sin t \\ \cos t \end{pmatrix}$$

$$\text{et } X_3(t) = \operatorname{Im} Z(t) = e^t \begin{pmatrix} 2\cos t \\ \cos t \\ \sin t \end{pmatrix}$$

On cherche ensuite une solution particulière pour les seconds membres $B_1(t)$ et $B_2(t)$ avec

$$B_1(t) = \begin{pmatrix} 0 \\ -2t \\ 0 \end{pmatrix} \text{ et } B_2(t) = \begin{pmatrix} 0 \\ 0 \\ -5\sin t \end{pmatrix}$$

et on additionnera ces deux solutions (principe de "superposition des solutions", qui ne fait que traduire la linéarité du système étudié). Pour B_1, on cherche cette solution sous la forme

$$Z_1(t) = \begin{pmatrix} at^2 + b \\ a't^2 + b' \\ a"t^2 + b" \end{pmatrix}$$

puisque 0 est valeur propre simple de A. On trouve

$$Z_1(t) = \begin{pmatrix} 6t^2 + 6t \\ 2t^2 + 3t \\ -3t - 3 \end{pmatrix}$$

Pour le second membre B_2, puisque la matrice A est réelle, on cherchera plutôt une solution particulière pour le second membre $\begin{pmatrix} 0 \\ 0 \\ -5e^{it} \end{pmatrix}$, et on en retiendra

la partie imaginaire. Cette solution est donc cherchée sous le forme

$$Z_2(t) = \operatorname{Im} Z(t) \text{ avec } Z(t) = e^{it}\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix}$$

puisque i n'est pas valeur propre de A. On résout

$$\begin{cases} i\alpha = -2\alpha + 6\beta - 2\gamma \\ i\beta = -\alpha + 3\beta - \gamma \\ i\gamma = -\alpha + 3\beta + \gamma - 5 \end{cases}$$

ce qui donne $\alpha = 2+4i, \beta = 1+2i, \gamma = 3+i$ et

$$Z_2(t) = \operatorname{Im} e^{it}\begin{pmatrix}2+4i\\ 1+2i\\ 3+i\end{pmatrix} = \begin{pmatrix}4\cos t + 2\sin t\\ \sin t + 2\cos t\\ 3\sin t + \cos t\end{pmatrix}$$

et on obtient finalement

$$\begin{cases} x(t) = 3\lambda - 2\mu e^t \sin t + 2\nu e^t \cos t + 6t^2 + 6t + 4\cos t + 2\sin t \\ y(t) = \lambda - \mu e^t \sin t + \nu e^t \cos t + 2t^2 + 3t + 2\cos t + \sin t \\ z(t) = \mu e^t \sin t + \nu e^t \sin t - 3t - 3 + \cos t + 3\sin t \end{cases}$$

avec $(\lambda, \mu, \nu) \in \mathbb{R}^3$ arbitraire.

19-1.2.5 Remarque

Pour un système homogène général

$$(E) : X'(t) = A(t)X(t)$$

à coefficients non constants, il n'existe pas d'expression simple donnant un système fondamental de solutions. En particulier, il ne faut pas croire que, pour $X_0 \in \mathbb{C}^n$, la fonction $t \mapsto \exp\left(\int_{t_0}^t A(u)\,du\right)X_0$ soit solution de (E). Ceci est dû au fait qu'en général, pour $B \in C^1(I, \mathcal{M}_n(\mathbb{C}))$, la dérivée de $t \mapsto \exp(B(t))$ n'est pas[10] $B'(t)\exp(B(t))$.
Par exemple avec $n = 2$ et

$$A(t) = \begin{pmatrix}0 & 2t\\ 1 & 0\end{pmatrix}$$

nous avons

$$\int_0^t A(u)\,du = \begin{pmatrix}0 & t^2\\ t & 0\end{pmatrix} = B(t)$$

[10] La démonstration faite dans le cas $n = 1$ n'est plus valable, notamment parce qu'en général

$$\exp(B(t+h)) \neq \exp(B(t+h) - B(t)) \times \exp B(t)$$

On vérifie facilement que

$$B^2(t) = t^3 I_2$$

ce qui donne

$$\forall n \in \mathbb{N} \quad B^{2n}(t) = t^{3n} I_2 \quad \text{et} \quad B^{2n+1}(t) = t^{3n} B(t)$$

et donc

$$\exp(B(t)) = \sum_{k=0}^{+\infty} \frac{B(t)^k}{k!} = \sum_{n=0}^{+\infty} \frac{t^{3n}}{(2n)!} I_2 + \sum_{n=0}^{+\infty} \frac{t^{3n}}{(2n+1)!} B(t)$$

En particulier, pour $t > 0$

$$\exp(B(t)) = \operatorname{ch}\left(t^{\frac{3}{2}}\right) I_2 + t^{-\frac{3}{2}} \operatorname{sh}\left(t^{\frac{3}{2}}\right) B(t) = \begin{pmatrix} \operatorname{ch}\left(t^{\frac{3}{2}}\right) & t^{\frac{1}{2}} \operatorname{sh}\left(t^{\frac{3}{2}}\right) \\ t^{-\frac{1}{2}} \operatorname{sh}\left(t^{\frac{3}{2}}\right) & \operatorname{ch}\left(t^{\frac{3}{2}}\right) \end{pmatrix}$$

On vérifie aisément que

$$X(t) = \exp(B(t)) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \operatorname{ch}\left(t^{\frac{3}{2}}\right) \\ t^{-\frac{1}{2}} \operatorname{sh}\left(t^{\frac{3}{2}}\right) \end{pmatrix}$$

n'est pas solution de $X'(t) = A(t) X(t)$ sur l'intervalle $]0, +\infty[$.

On peut cependant trouver la solution de ce système pour la condition initiale

$$X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

en utilisant un développement en série entière. Si $X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ le système $X'(t) = A(t) X(t)$ s'écrit

$$\begin{cases} x'(t) = 2t\, y(t) \\ y'(t) = x(t) \end{cases}$$

Pour une solution donnée, la fonction $t \mapsto y(t)$ est classe $\mathcal{C}^2$, puisque sa dérivée est de classe $\mathcal{C}^1$ (on montrerait d'ailleurs par une récurrence simple que x et y sont de classe $\mathcal{C}^\infty$), avec $y''(t) = x'(t)$, soit

$$(E_2) : y''(t) = 2t\, y(t)$$

On cherche donc une solution de cette équation (du second ordre) qui vérifie $y(0) = 0$ et $y'(0) = x(0) = 1$ sous forme d'une série entière (*cf.* section **11-3.3.5**), en raisonnant par analyse et synthèse.

Analyse : si la somme d'une série entière

$$f(t) = \sum_{n=0}^{+\infty} a_n t^n$$

de rayon de convergence $R > 0$ vérifie l'équation (E_2) dans l'intervalle $]-R, R[$, on a

$$(*) : \forall t \in]-R, R[\quad \sum_{n=0}^{+\infty} (n+2)(n+1) a_{n+2} t^n = 2 \sum_{n=0}^{+\infty} a_n t^{n+1}$$

ce qui donne, par unicité du développement en série entière

$$(**) : \begin{cases} a_2 = 0 \\ \forall n \geqslant 1 \quad (n+2)(n+1) a_{n+2} = 2a_{n-1} \end{cases}$$

Comme on cherche une solution avec $y(0) = 0$ et $y'(0) = 1$, on rajoute les conditions $a_0 = 0$ et $a_1 = 1$. La relation de récurrence précédente montre alors que, pour tout $n \geqslant 0$, $a_{3n} = a_{3n+2} = 0$, alors que

$$\forall n \geqslant 1 \quad a_{3n+1} = \frac{2}{(3n+1)(3n+2)} a_{3n-2} = \frac{6(n+1)}{(3n+1)(3n+2)(3n+3)} a_{3n-2}$$

ce qui donne finalement

$$a_{3n+1} = \frac{6^{n+1}(n+1)!}{(3n+3)!} a_1 = \frac{6^{n+1}(n+1)!}{(3n+3)!}$$

et donc

$$f(t) = \sum_{n=0}^{+\infty} \frac{6^{n+1}(n+1)!}{(3n+3)!} t^{3n+1}$$

Synthèse : la série entière précédente a un rayon de convergence infini et sa somme est solution sur $\mathbf{R}$ de (E_2) : en effet, la suite de ses coefficients $(a_n)_{n \in \mathbf{N}}$ vérifie la relation $(**)$, ce qui donne (règle de D'Alembert) $R = +\infty$ et la relation $(*)$. Si on pose

$$y(t) = \sum_{n=0}^{+\infty} \frac{6^{n+1}(n+1)!}{(3n+3)!} t^{3n+1} \quad \text{et} \quad x(t) = y'(t) = \sum_{n=0}^{+\infty} \frac{6^{n+1}(3n+1)(n+1)!}{(3n+3)!} t^{3n}$$

on vérifie ainsi que $X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ est solution du système pour la condition initiale

$$X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

19-1.3 Equations scalaires d'ordre 2

L'exemple qui précède montre qu'une équation différentielle (linéaire) du second ordre peut être reliée à un système différentiel du premier ordre. Cette situation est générale, et nous verrons dans cette section comment la théorie des équations linéaires étudiée précédemment englobe celle des équations d'ordre plus élevé. Nous étudions d'abord le cas des équations scalaires du second ordre.

DÉFINITION 19-1.33 *On appelle équation différentielle linéaire scalaire du second ordre toute équation de la forme*

$$(E) : x''(t) + a(t)\, x'(t) + b(t)\, x(t) = c(t)$$

où I un intervalle de $\mathbb{R}$ et $a, b, c \in C^0(I, \mathbb{K})$. Une solution de (E) est une fonction $f : I \to \mathbb{K}$ deux fois dérivable[11] *vérifiant*

$$\forall t \in I \quad f''(t) + a(t)\, f'(t) + b(t)\, f(t) = c(t)$$

L'équation homogène associée à (E) est l'équation

$$(E_H) : x''(t) + a(t)\, x'(t) + b(t)\, x(t) = 0$$

19-1.3.1 Système du premier ordre équivalent

PROPOSITION 19-1.34 Une fonction de classe C^2 $x : I \to \mathbb{K}$ est solution de (E) si et seulement si la fonction (de classe C^1) $X : I \to \mathbb{K}^2$ définie par $X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}$ est solution du système

$$(\Sigma_E) : X'(t) = A(t)\, X(t) + B(t)$$

où $A : I \to \mathcal{M}_2(\mathbb{K})$ et $B : I \to \mathbb{K}^2$ sont définies par

$$A(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$$

Réciproquement, si $X : I \to \mathbb{K}^2$ est une fonction de classe C^1 solution de (Σ_E) avec

$$\forall t \in I\ X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

la fonction $x_1 : I \to \mathbb{K}$ est de classe C^2 et est solution de (E) sur I (et x_2 est égale à la dérivée de x_1). Le système (Σ_E) est donc appelé système linéaire du premier ordre équivalent à l'équation (E).

Démonstration : Si $x \in C^2(I, \mathbb{K})$ est solution de (E), on a

$$\forall t \in I \quad \begin{pmatrix} x'(t) \\ x''(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} + \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$$

et donc X est solution de (Σ_E). Réciproquement, si

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in C^1(I, \mathbb{K})$$

est solution de (Σ_E), on a

$$\forall t \in I \quad \begin{cases} x_1'(t) = x_2(t) \\ x_2'(t) + a(t)\, x_2(t) + b(t)\, x_1(t) = c(t) \end{cases}$$

[11] En fait de classe C^2

donc $x_1' \in \mathcal{C}^1(I,\mathbb{K})$, ce qui prouve $x_1 \in \mathcal{C}^2(I,\mathbb{K})$ et

$$\forall t \in I \quad x_1''(t) + a(t)\,x_2(t) + b(t)\,x_1(t) = c(t)$$

et x_1 est bien solution de (E) sur I. ■

La proposition qui précède montre simplement que l'application linéaire φ : $\mathcal{C}^2(I,\mathbb{K}) \to \mathcal{C}^1(I,\mathbb{K}^2)$ définie par

$$\varphi : x \mapsto X = \begin{pmatrix} x \\ x' \end{pmatrix}$$

induit une bijection entre l'ensemble des solutions de (E) et l'ensemble des solutions de (Σ_E). Le théorème de Cauchy appliqué à (Σ_E) donne donc :

THÉORÈME 19-1.35 Si $t_0 \in I$ et $(x_0, x_1) \in \mathbb{K}^2$, l'équation

$$(E) : x''(t) + a(t)\,x'(t) + b(t)\,x(t) = c(t)$$

où a, b et $c \in \mathcal{C}^0(I,\mathbb{K})$ possède une unique solution, définie sur I, vérifiant la condition initiale

$$\begin{cases} x(t_0) = x_0 \\ x'(t_0) = x_1 \end{cases}$$

19-1.3.2 Solutions de l'équation homogène

Le système du premier ordre équivalent à (E_H) est lui-même homogène. L'ensemble de ses solutions est un $\mathbb{K}$-espace vectoriel de dimension 2. On en déduit :

PROPOSITION 19-1.36 L'ensemble $\mathcal{S}_{E_H}$ des solutions de (E_H) est un $\mathbb{K}$-ev de dimension 2. Si $t_0 \in I$ est fixé, l'application

$$\mathcal{S}_{E_H} \to \mathbb{K}^2 \quad x \mapsto \begin{pmatrix} x(t_0) \\ x'(t_0) \end{pmatrix}$$

est un isomorphisme d'espaces vectoriels.

DÉFINITION 19-1.37 *Un système fondamental de solutions de (E_H) est une base de l'espace $\mathcal{S}_{E_H}$.*

Deux solutions x_1 et x_2 de (E_H) forment un système fondamental de solutions si et seulement si on peut trouver un point $t_0 \in I$ tel que les vecteurs de $\mathbb{K}^2$ correspondant aux "conditions initiales"

$$X_1(t_0) = \begin{pmatrix} x_1(t_0) \\ x_1'(t_0) \end{pmatrix} \quad \text{et} \quad X_2(t_0) = \begin{pmatrix} x_2(t_0) \\ x_2'(t_0) \end{pmatrix}$$

soient indépendants. Lorsque cette propriété est vérifiée en t_0, elle l'est en tout point $t \in I$.

DÉFINITION 19-1.38 *Si x_1 et x_2 sont deux fonctions de $\mathcal{C}^1(I,\mathbb{K})$, on appelle wronskien de ces deux fonctions l'application*

$$W_{x_1,x_2} : I \to \mathbb{K} \quad t \mapsto W_{x_1,x_2}(t) = \begin{vmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{vmatrix} = x_1(t)\,x_2'(t) - x_1'(t)\,x_2(t)$$

PROPOSITION 19-1.39 Si x_1 et x_2 sont dans $\mathcal{S}_{E_H}$, on a deux possibilités qui s'excluent mutuellement :

- **soit la fonction W_{x_1,x_2} est identiquement nulle sur I, et (x_1, x_2) est une famille liée.**
- **soit la fonction W_{x_1,x_2} ne s'annule pas sur I, et (x_1, x_2) est un système fondamental de solutions de (E_H).**

Toute solution $x \in \Sigma_{E_H}$ s'écrit alors

$$x = \lambda_1 x_1 + \lambda_2 x_2$$

où λ_1 et $\lambda_2 \in \mathbb{K}$ sont deux constantes arbitraires.

La formule de Liouville (qu'on pourrait obtenir en appliquant la proposition **19-1**.12 aux solutions X_1 et X_2 de Σ_{E_H} correspondant à x_1 et x_2) a ici une démonstration très simple : si x_1 et x_2 sont deux solutions de (E_H), on a pour tout $t \in I$

$$\begin{aligned} W'_{x_1,x_2}(t) &= x_1(t)\, x''_2(t) - x''_1(t)\, x_2(t) \\ &= x_1(t)\left[-a(t)\, x'_2(t) - b(t)\, x_2(t)\right] - x_2(t)\left[-a(t)\, x'_1(t) - b(t)\, x_1(t)\right] \end{aligned}$$

ce qui donne immédiatement $\forall t \in I \quad W'_{x_1,x_2}(t) = -a(t)\, W_{x_1,x_2}(t)$ et donc, pour $t_0 \in I$ arbitraire

$$\boxed{\forall t \in I \quad W_{x_1,x_2}(t) = W_{x_1,x_2}(t_0)\, e^{-\int_{t_0}^{t} a(u)\, du}}$$

On remarquera qu'en particulier, lorsque la fonction a est identiquement nulle, le wronskien de deux solutions de (E_H) est constant.

REMARQUE 19-1.40 Pour une équation de la forme

$$\alpha(t)\; x''(t) + \beta(t)\; x'(t) + \gamma(t)\; x(t) = \delta(t)$$

où α, β, γ et $\delta \in C^0(I, \mathbb{K})$, les résultats qui précèdent s'appliquent sur tout sous-intervalle $J \subset I$ sur lequel la fonction α ne s'annule pas (voir remarque **19-1**.18).

Il n'y a pas de méthode générale pour expliciter un système fondamental de solutions de (E_H). A part le cas des équations à coefficients constants (voir section **19-1.3.5**), on pourra cependant résoudre explicitement certaines équations en utilisant les remarques qui suivent :

REMARQUE 19-1.41 Si on connaît une solution particulière de (E_H) qui ne s'annule pas sur I :

Si u est cette solution, on cherchera les solutions de (E_H) en faisant le changement de fonction inconnue

$$x = uz$$

ce qui revient à considérer la fonction $z = \frac{x}{u}$, parfaitement définie et de classe C^2 si x l'est, puisque le dénominateur ne s'annule pas. On a alors

$$x' = u'z + uz' \quad \text{et} \quad x'' = u''z + 2u'z' + uz''$$

En reportant dans (E_H), on obtient

$$u\,z'' + (2u' + au)\ z' + (u'' + au' + bu)\ z = 0$$

et le terme en z s'annule, puisque u est solution de (E_H). La fonction z est donc solution de

$$u\,z'' + (2u' + au)\ z' = 0$$

qui est une équation linéaire du premier ordre vérifiée par la fonction z'. On peut donc expliciter z' par la méthode vue à la section **19-1.1.5**, et en déduire z par intégration.

EXERCICE 19-1.42 Trouver un système fondamental de solutions sur $]0, +\infty[$ de l'équation

$$t^3\,x''(t) - t\,x'(t) + x(t) = 0$$

en remarquant que $u(t) = t$ est une solution particulière. (Réponse : $x_1(t) = u(t) = t$ et $x_2(t) = t\,e^{-\frac{1}{t}}$).

REMARQUE 19-1.43 Equations de la forme

$$\alpha(t)\,x''(t) + \beta(t)(t)\ x'(t) + \gamma(t)\ x(t) = 0$$

où α, β et γ sont des fonctions polynômes (ou parfois des fonctions développables en série entière dans un intervalle $]-R, R[$ où l'on cherche à résoudre l'équation). On peut parfois trouver des solutions sous forme de séries entières

$$x(t) = \sum_{n=0}^{+\infty} a_n\,t^n$$

en raisonnant par analyse / synthèse, comme nous l'avons fait à la section **19-1.2.5** pour l'équation

$$y''(t) - 2t\,y(t) = 0$$

EXERCICE 19-1.44 Trouver un système fondamental de solutions sur $]0, +\infty[$ de l'équation

$$t\,x''(t) - x'(t) + 4t^3\,x(t) = 0$$

Expliciter ces solutions à l'aide des fonctions "usuelles". Montrer que l'ensemble $\mathcal{S}$ des fonctions de classe C^2 solutions de cette équation sur $\mathbb{R}$ est un espace vectoriel de dimension 2, mais que l'application

$$\mathcal{S} \to \mathbb{R}^2 \quad x \mapsto (x(0), x'(0))$$

n'est pas un isomorphisme d'espaces vectoriels.

EXERCICE 19-1.45 Résoudre l'équation

$$t^2(1-t)\,x''(t) - t(1+t)\,x'(t) + x(t) = 0$$

En cherchant une solution sous forme de série entière, on trouvera une fraction rationnelle solution dans $]-1,1[$, mais aussi dans $]-\infty,1[$ ou $]1,+\infty[$. On utilisera ensuite la méthode décrite à la remarque **19-1**.41. Montrer que sur $]-\infty,1[$ et sur $]0,+\infty[$, l'ensemble des fonctions de classe C^2 vérifiant cette équation sur $]-\infty,1[$ (ou sur $]0,+\infty[$) est un espace de dimension 1. (Réponse : sur $]-\infty,1[$, c'est l'espace engendré par $x \mapsto \dfrac{x}{1-x}$; sur $]0,+\infty[$, on trouvera la fonction $x \mapsto \dfrac{\ln x}{1-x}$ convenablement prolongée en 1).

EXERCICE 19-1.46 Soit $q :]-R,R[\to \mathbb{K}$ la somme d'une série entière. Montrer que toutes les solutions dans $]-R,R[$ de l'équation différentielle

$$x''(t) - q(t)\,x(t) = 0$$

sont développables en série entière.

REMARQUE 19-1.47 Utilisation d'un changement de variable.

Pour certaines équations (construites sur mesure), on peut parfois chercher des solutions sous la forme

$$x(t) = z(\varphi(t))$$

où $\varphi : I \to J = \varphi(I)$ est une fonction de classe C^2 et $z : J \to \mathbb{K}$ est la nouvelle fonction inconnue supposée elle-aussi de classe C^2. Il faut alors choisir φ pour que l'équation vérifiée par z soit "simple". Par exemple, pour l'équation

$$t\,x''(t) - x'(t) + 4t^3\,x(t) = 0$$

vue à l'exercice **19-1**.44, en posant $x(t) = z(\varphi(t))$, on a

$$x'(t) = \varphi'(t)\,z'(\varphi(t)) \text{ et } x''(t) = [\varphi'(t)]^2\,z''(\varphi(t)) + \varphi''(t)\,z'(\varphi(t))$$

ce qui donne, en reportant dans l'équation

$$t\,[\varphi'(t)]^2\,z''(\varphi(t)) + (t\varphi''(t) - \varphi'(t))\,z'(\varphi(t)) + 4t^3\,z(\varphi(t)) = 0$$

En prenant $\varphi(t) = t^2$, on retombe sur une équation à coefficients constants.

EXERCICE 19-1.48 Utiliser cette méthode pour résoudre

$$(1+t^2)\,x''(t) + t\,x'(t) - \frac{x(t)}{4} = 0$$

(on pourrait ici aussi rechercher des solutions sous forme de séries entières).

19-1.3.3 Méthode de variation des constantes

On suppose connu un système fondamental $(x_i)_{i=1,2}$ de solutions de

$$(E_H) : x''(t) + a(t)\, x'(t) + b(t)\, x(t) = 0$$

et on cherche à résoudre

$$(E) : x''(t) + a(t)\, x'(t) + b(t)\, x(t) = c(t)$$

On applique la méthode de variations des constantes au système du premier ordre équivalent Σ_E, c'est à dire que l'on cherche le vecteur $X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}$ sous la forme

$$\begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} = C_1(t) \begin{pmatrix} x_1(t) \\ x_1'(t) \end{pmatrix} + C_2(t) \begin{pmatrix} x_2(t) \\ x_2'(t) \end{pmatrix}$$

où C_1 et C_2 sont deux fonctions inconnues de classe C^1 sur I. En reportant dans le système Σ_E (*cf.* proposition **19-1**.34), on obtient

$$C_1'(t) \begin{pmatrix} x_1(t) \\ x_1'(t) \end{pmatrix} + C_2'(t) \begin{pmatrix} x_2(t) \\ x_2'(t) \end{pmatrix} = B(t) = \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$$

ce qui permet de déterminer C_1' et C_2' en résolvant un système de Cramer, puisque le déterminant

$$\begin{vmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{vmatrix} = W_{x_1,x_2}(t)$$

ne s'annule jamais. On en déduit les fonctions C_1 et C_2, puis

$$x(t) = C_1(t)\, x_1(t) + C_2(t)\, x_2(t)$$

EXERCICE 19-1.49 (suite de l'exercice **19-1**.42) Résoudre dans l'intervalle $]0, +\infty[$

$$t^3 x''(t) - t\, x'(t) + x(t) = \sin\left(\frac{1}{t}\right)$$

Attention ! Le second membre vaut

$$c(t) = \frac{1}{t^3} \sin\left(\frac{1}{t}\right)$$

Réponse :

$$x(t) = -\frac{t}{2}\left(\cos\left(\frac{1}{t}\right) + \sin\left(\frac{1}{t}\right)\right) + \lambda t + \mu t\, e^{-\frac{1}{t}}$$

où λ et μ sont deux constantes réelles.

Bien entendu, il est parfois possible d'éviter la méthode de variation des constantes, en cherchant par d'autres méthodes une solution particulière de l'équation avec second membre, et en utilisant la structure d'espace affine de l'ensemble des solutions. Voir à cet effet l'exercice **19-1**.54.

19-1.3.4 Remarque

Nous nous sommes limités au cas des équations scalaires d'ordre 2. Dans le cas d'une équation linéaire scalaire d'ordre n

$$(E_n) : x^{(n)}(t) + a_1(t)\, x^{(n-1)}(t) + \cdots + a_{n-1}(t)\, x'(t) + a_n(t)\, x(t) = c(t)$$

avec $a_i \in \mathcal{C}^0(I, \mathbb{K})$ pour $i = 1, \ldots, n$ et $c \in \mathcal{C}^0(I, \mathbb{K})$, on obtiendrait de même un système du premier ordre équivalent

$$X'(t) = A(t)\, X(t) + B(t)$$

avec

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_n(t) & -a_{n-1}(t) & \cdots & -a_2(t) & -a_1(t) \end{pmatrix}$$

$$X(t) = \begin{pmatrix} x(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} \text{ et } B(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c(t) \end{pmatrix}$$

La théorie se développe de la même façon : l'ensemble des solutions de l'équation sans second membre est un espace vectoriel de dimension n. Un système $(x_i)_{1 \leqslant i \leqslant n}$ de solutions de cette équation est une base de cet espace (un système fondamental de solutions) si et seulement si

$$\forall t \in I \quad W(t) = \begin{vmatrix} x_1(t) & \cdots & x_n(t) \\ \vdots & \vdots & \vdots \\ x_1^{(n-1)}(t) & \cdots & x_n^{(n-1)}(t) \end{vmatrix} \neq 0$$

On a d'ailleurs, pour $t_0 \in I$ arbitraire

$$\forall t \in I \quad W(t) = W(t_0)\, e^{-\int_{t_0}^{t} a_1(u)\, du}$$

Si on connaît un tel système fondamental, une solution de (E_n) pourra s'écrire

$$x(t) = \sum_{i=1}^{n} C_i(t)\, x_i(t)$$

où les fonctions $(C_i)_{1 \leqslant i \leqslant n}$ de classe $\mathcal{C}^1$ vérifient

$$C_1'(t) \begin{pmatrix} x_1(t) \\ \vdots \\ x_1^{(n-1)}(t) \end{pmatrix} + \cdots + C_n'(t) \begin{pmatrix} x_n(t) \\ \vdots \\ x_n^{(n-1)}(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c(t) \end{pmatrix}$$

Il est à noter que, dans ce cas, une "condition initiale" correspond au choix de $t_0 \in I$ et $(\alpha_0, \ldots, \alpha_{n-1}) \in \mathbb{K}^n$, et on a existence et unicité d'une solution vérifiant

$$\forall i \in \{0, \ldots, n-1\} \quad x^{(i)}(t_0) = \alpha_i$$

19-1.3.5 Equations à coefficients constants

Nous étudions ici les équations de la forme

$$(E) : x^{(n)}(t) + a_1 x^{(n-1)}(t) + \cdots + a_{n-1} x'(t) + a_n x(t) = c(t)$$

où $(a_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n$ et $c \in \mathcal{C}^0(I, \mathbb{K})$. Nous cherchons dans ce cas d'abord à expliciter une base du $\mathbb{K}$-ev de dimension n constitué des solutions de l'équation sans second membre. Nous nous limiterons au cas où $\mathbb{K} = \mathbb{C}$. Lorsque les $(a_i)_{1 \leqslant i \leqslant n}$ sont réels, on obtiendra facilement un système fondamental de solutions réelles à l'aide des parties réelles et imaginaires des solutions complexes. On notera l'analogie avec l'étude de la section **5-7**.

On pourrait bien sûr étudier le système linéaire du premier ordre équivalent à l'équation

$$(E_H) : x^{(n)}(t) + a_1 x^{(n-1)}(t) + \cdots + a_{n-1} x'(t) + a_n x(t) = 0$$

La matrice A de ce système

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

est à coefficients constants, et on peut écrire les solutions de (E_H) en utilisant la matrice $\exp(tA)$. Il est plus simple ici d'utiliser le théorème de décomposition des noyaux : on remarque d'abord que toutes les solutions de (E_H) sont dans $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$. Si on note D l'endomorphisme de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ correspondant à la dérivation,

$$D : x \mapsto x'$$

on voit que l'équation (E_H) peut s'écrire

$$\left[D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right](x) = 0_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}$$

et donc

$$P(D)(x) = 0_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}$$

L'espace $\mathcal{S}_{E_H}$ des solutions est donc simplement

$$\mathcal{S}_{E_H} = \ker P(D)$$

Pour obtenir une décomposition de ce noyau, on factorise le polynôme P en facteurs irréductibles dans $\mathbb{C}[X]$

$$P(X) = X^n + a_1 X^{n-1} + \cdots + a_n = \prod_{i=1}^{p} (X - r_i)^{n_i}$$

où les $(r_i)_{1\leqslant i\leqslant p}$ sont les racines distinctes de l'**équation caractéristique**

$$P(r) = r^n + a_1 r^{n-1} + \cdots + a_n = 0$$

de multiplicités respectives (n_i). Ces racines correspondent aux nombres complexes r tels que $t \mapsto e^{rt}$ soit solution de (E_H), puisque, pour $x(t) = e^{rt}$ on a évidemment

$$[P(D)(x)](t) = P(r)\, e^{rt}$$

Comme pour $i \neq j$ on a $(X - r_i)^{n_i} \wedge (X - r_j)^{n_j} = 1$, le théorème de décomposition des noyaux donne

$$\mathcal{S}_{E_H} = \ker P(D) = \bigoplus_{i=1}^{p} \ker \left(D - r_i\, Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right)^{n_i}$$

Chacun des sous-espaces $\ker\left(D - r_i\, Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right)^{n_i}$ est de dimension n_i, puisque c'est l'ensemble des solutions d'une équation différentielle linéaire d'ordre n_i, sans second membre (et à coefficients constants). Il suffit d'en trouver n_i éléments indépendants pour en avoir une base. Par recollement, on obtiendra ensuite une base de $\mathcal{S}_{E_H}$. Si $n_i = 1$, on a évidemment

$$\ker\left(D - r_i\, Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right) = \mathrm{vect}\left(t \mapsto e^{r_i t}\right)$$

Dans le cas où $n_i > 1$, on remarque que, pour $Q \in \mathbb{C}[X]$ et $x : \mathbb{R} \to \mathbb{C}$ définie par

$$x(t) = Q(t)\, e^{r_i t}$$

on a

$$x'(t) - r_i x(t) = Q'(t)\, e^{r_i t}$$

ce qui donne en itérant

$$\left[\left(D - r_i\, Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right)^{n_i}\right](x)(t) = Q^{(n_i)}(t)\, e^{r_i t}$$

et donc cette fonction est nulle si Q est de degré inférieur ou égal à $n_i - 1$. On en déduit facilement que la famille de fonctions $\left(t \mapsto t^k e^{r_i t}\right)_{0\leqslant k\leqslant n_i-1}$ est une base de $\ker\left(D - r_i\, Id_{\mathcal{C}^\infty(\mathbb{R},\mathbb{C})}\right)^{n_i}$. Les solutions de (E_H) sont donc exactement les combinaisons linéaires des fonctions

$$\left(t \mapsto t^k e^{r_i t}\right)_{\substack{0\leqslant k\leqslant n_i-1\\ 1\leqslant i\leqslant p}}$$

Résumons l'étude qui précède :

PROPOSITION 19-1.50 Les solutions de l'équation à coefficients constants (où les $(a_i)_{1\leqslant i\leqslant n}$ sont des complexes)

$$x^{(n)} + a_1 x^{(n-1)} + \cdots + a_n x = 0$$

sont les fonctions qui peuvent s'écrire

$$x(t) = \sum_{i=1}^{p} Q_i(t)\, e^{r_i t}$$

avec $Q_i \in \mathbb{C}_{n_i-1}[X]$, où les $(r_i)_{1\leqslant i\leqslant p}$ sont les racines distinctes de l'équation caractéristique

$$r^n + a_1 r^{n-1} + \cdots + a_n = 0$$

et les $(n_i)_{1\leqslant i\leqslant p}$ sont leurs multiplicités.

En particulier pour l'équation différentielle

$$x'' + ax' + bx = 0$$

du second ordre à coefficients constants, on obtient :

- **Si l'équation caractéristique a deux racines distinctes** r_1 et r_2, les solutions sont de la forme

$$x(t) = \lambda\, e^{r_1 t} + \mu\, e^{r_2 t} \quad \text{avec} \quad \lambda, \mu \in \mathbb{C}$$

- **Si l'équation caractéristique a une racine double** $r_1 = r_2 = r$, les solutions sont

$$x(t) = (\lambda\, t + \mu)\; e^{rt}$$

Lorsque a et $b \in \mathbb{R}$ avec $a^2 - 4b < 0$, les racines r_1 et r_2 sont complexes conjuguées

$$r_1 = \alpha + i\omega \quad \text{et} \quad r_2 = \alpha - i\omega \quad \text{avec } (\alpha, \omega) \in \mathbb{R} \times \mathbb{R}^*$$

un système fondamental de solutions réelles sera donné par

$$x_1(t) = \mathrm{Re}\left(e^{r_1 t}\right) = e^{\alpha t} \cos \omega t \quad \text{et} \quad x_2(t) = \mathrm{Im}\left(e^{r_1 t}\right) = e^{\alpha t} \sin \omega t$$

Pour résoudre ensuite l'équation avec second membre

$$x^{(n)}(t) + a_1\, x^{(n-1)}(t) + \cdots + a_{n-1}\, x'(t) + a_n\, x(t) = c(t)$$

on pourra utiliser la méthode de variation des constantes décrite dans les sections **19-1.3.3** et **19-1.3.4** :

EXERCICE 19-1.51 Résoudre sur $]0, +\infty[$

$$x''(t) + 3x'(t) + 2x(t) = \frac{t-1}{t^2} e^{-t}$$

Lorsque le second membre est de la forme

$$c(t) = e^{\mu t} Q(t)$$

où Q est un polynôme, on préfère souvent chercher une solution particulière par une méthode de coefficients indéterminés, en utilisant la proposition :

PROPOSITION 19-1.52 Si $Q \in \mathbb{K}[X]$ est un polynôme de degré m et $\mu \in \mathbb{K}$, l'équation à coefficients constants

$$x^{(n)}(t) + a_1 x^{(n-1)}(t) + \cdots + a_{n-1} x'(t) + a_n x(t) = e^{\mu t} Q(t)$$

(avec$(a_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n$) possède une solution particulière de la forme

$$x(t) = e^{\mu t} R(t)$$

où $R \in \mathbb{K}[X]$ est un polynôme de degré m si μ n'est pas racine de l'équation caractéristique et de degré $m+q$ si μ est racine d'ordre q de cette équation.

Démonstration : (Indication) : en utilisant la factorisation du polynôme

$$P(X) = X^n + a_1 X^{n-1} + \cdots + a_n = \prod_{i=1}^{p} (X - r_i)^{n_i}$$

montrer que, pour $R \in \mathbb{K}[X]$ et pour x définie par $x(t) = e^{\mu t} R(t)$, la fonction $y = P(D)(x)$ est de la forme

$$y(t) = e^{\mu t} R_1(t)$$

où R_1 est un polynôme de $\mathbb{K}[X]$. Montrer que, si le degré de R est inférieur à $m+q$, alors le degré de R_1 est inférieur à m. Montrer que la correspondance $R \mapsto R_1$ définit un morphisme de $\mathbb{K}_{m+q}[X]$ dans $\mathbb{K}_m[X]$. Quel est son noyau ? Conclure. ■

EXERCICE 19-1.53 Trouver les solutions réelles de

$$x'''(t) - x(t) = e^t \cos(t)$$

On cherchera une solution pour le second membre $c(t) = e^{(1+i)t}$, dont on conservera la partie réelle. Réponse :

$$x(t) = \lambda e^t + \mu e^{-\frac{t}{2}} \cos\left(\frac{\sqrt{3}}{2} t\right) + \nu e^{-\frac{t}{2}} \sin\left(\frac{\sqrt{3}}{2} t\right) + \frac{e^{-t}}{13} (3\cos t - 2\sin t)$$

avec λ, μ et $\nu \in \mathbb{R}$.

EXERCICE 19-1.54 En utilisant la remarque **19-1**.47, résoudre

$$tx''(t) - x'(t) + 4t^3 x(t) = 8t^3 e^{t^2}$$

EXERCICE 19-1.55 Résoudre dans $]0,+\infty[$ l'équation

$$t^3x'''(t)+t^2x''(t)-2tx'(t)+2x(t)=t^3e^t$$

Il s'agit d'une équation dont les coefficients sont des monômes en t de degrés "échelonnés". On peut donc chercher des fonctions solutions de l'équation homogène de la forme $t\mapsto t^\alpha$ et on tombe sur une équation du troisième degré en α. Une autre présentation du même calcul consiste à faire un changement de variable en cherchant x solution de l'équation homogène sous la forme

$$x(t)=z(\ln(t))$$

avec $z\in\mathcal{C}^3(\mathbb{R},\mathbb{C})$.

19-2 Equations non linéaires

19-2.1 Equations autonomes

19-2.1.1 Définitions

Pour simplifier, nous nous placerons sur l'espace vectoriel $\mathbb{R}^p$, muni d'une norme quelconque. On appelle équation différentielle autonome du premier ordre toute équation[12] de la forme

$$(E)\quad x'(t)=f(x(t))$$

où f est une application de classe $\mathcal{C}^1$ d'un ouvert $\mathcal{U}$ de $\mathbb{R}^p$ dans $\mathbb{R}^p$.

DÉFINITION 19-2.1 *Si I est un intervalle de $\mathbb{R}$ non réduit à un point, une I-solution de (E) est une application dérivable*[13]

$$z:I\to\mathcal{U}$$

vérifiant

$$\forall t\in I\quad z'(t)=f(z(t))$$

L'équation (E) est dite "autonome" car le second membre ne dépend pas du temps, contrairement à ce qui se passe pour une équation de la forme

$$x'(t)=f(t,x(t))$$

Il en résulte évidemment qu'une translation sur le temps conserve les solutions de (E). Plus précisément :

---

[12] On la notera souvent simplement $x'=f(x)$

[13] Avec les hypothèses faites sur f, elle est en fait de classe $\mathcal{C}^2$. On notera qu'il est important d'avoir

$$\forall t\in I\quad z(t)\in\mathcal{U}$$

pour que cette équation ait un sens.

PROPOSITION 19-2.2 Si I est un intervalle de $\mathbb{R}$ et z est une I-solution de (E), alors pour tout $t_0 \in \mathbb{R}$ la fonction y définie par

$$y : t_0 + I \to \mathcal{U} \quad t \mapsto y(t) = z(t - t_0)$$

est une $(t_0 + I)$-solution de (E).

Bien sûr, l'équation (E) formalise en fait un système de p équations "scalaires"

$$\begin{cases} x_1'(t) = f_1(x_1(t), \dots, x_p(t)) \\ \vdots \\ x_p'(t) = f_p(x_1(t), \dots, x_p(t)) \end{cases}$$

où les $(f_i)_{1 \leqslant i \leqslant p}$ sont p applications de classe C^1 de $\mathcal{U}$ dans $\mathbb{R}$, et où

$$x(t) = (x_1(t), \dots, x_p(t))$$

est la fonction inconnue.

DÉFINITION 19-2.3 *Une courbe intégrale de l'équation (E) est un arc paramétré (dont le support est inclus dans $\mathcal{U}$)*

$$I \ni t \mapsto x(t) \in \mathcal{U}.$$

où x est une I-solution de (E). Le support d'une telle courbe intégrale est aussi appelé trajectoire d'une solution de (E).

On peut interpréter l'équation (E) et ses courbes intégrales avec le vocabulaire des champs de vecteurs : l'application

$$f : \mathcal{U} \to \mathbb{R}^p \quad x \mapsto f(x)$$

peut être considérée comme la donnée d'un champ de vecteurs (stationnaire, c'est-à-dire indépendant de t) dans $\mathcal{U}$. Une courbe intégrale est un arc de classe C^1 tracé dans $\mathcal{U}$, $t \mapsto x(t)$ paramétré par $t \in I$. Si x_0 est un point du support de cet arc où le champ ne s'annule pas ($f(x_0) \neq 0$), cet arc est régulier pour toute valeur du paramètre $t_0 \in I$ telle que $x(t_0) = x_0$, et la tangente à l'arc pour cette valeur du paramètre est dirigée par le vecteur $x'(t_0) = f(x_0)$. Une trajectoire d'une solution de (E) est donc une ligne de champ de f. Nous verrons aussi ultérieurement qu'avec l'hypothèse "$f \in C^1(\mathcal{U}, \mathbb{R}^p)$", toute trajectoire contenant un point x_0 où f s'annule se réduit au point x_0. On donne alors la définition :

DÉFINITION 19-2.4 *Un point $x_0 \in \mathcal{U}$ est un équilibre pour l'équation différentielle $x' = f(x)$ si et seulement si $f(x_0) = 0$. Il est alors clair que la fonction constante égale à x_0 est une $\mathbb{R}$-solution de cette équation.*

Comme dans le cas des équations linéaires, une équation autonome du second ordre

$$(E_2) : x''(t) = g(x(t), x'(t))$$

(où g est une application de classe $\mathcal{C}^1$ d'un ouvert $\mathcal{U}$ de $\mathbb{R}^p \times \mathbb{R}^p$ dans $\mathbb{R}^p$) peut se ramener à une équation du premier ordre équivalente[14] : il suffit de considérer l'équation

$$(\Sigma_{E_2}) : z'(t) = f(z(t))$$

où f est l'application de $\mathcal{U}$ dans $\mathbb{R}^p \times \mathbb{R}^p$ définie par

$$\forall (x_1, x_2) \in \mathcal{U} \quad f(x_1, x_2) = (x_2, g(x_1, x_2))$$

évidemment de classe $\mathcal{C}^1$ si g l'est. On vérifie alors aisément que, si I est un intervalle de $\mathbb{R}$, une application de classe $\mathcal{C}^1$

$$z : I \to \mathbb{R}^p \times \mathbb{R}^p \quad t \mapsto z(t) = (z_1(t), z_2(t))$$

est une I-solution de (Σ_{E_2}) si et seulement si $z_1 \in \mathcal{C}^2(I, \mathbb{R}^p)$ est I-solution[15] de (E_2), avec $z_1' = z_2$: l'espace "naturel" pour décrire les solutions de (E_2) n'est pas un ouvert de $\mathbb{R}^p$, mais l'ouvert $\mathcal{U}$ de $\mathbb{R}^{2p}$. Nous verrons qu'avec les hypothèses faites sur g, c'est la donnée d'une condition initiale $(x(t_0), x'(t_0)) \in \mathcal{U}$ qui caractérise entièrement une solution x de (E_2). C'est donc la trajectoire de $t \mapsto (x(t), x'(t))$ qu'il faut suivre. On dit que l'espace "des phases" est l'ouvert $\mathcal{U}$ (de dimension $2p$).

Nous nous limiterons dans ce qui suit aux énoncés concernant les équations du premier ordre, ce qui n'est donc pas une restriction théorique.

19-2.1.2 Restriction d'une solution, recollement, solution maximale

On considère l'équation $(E) : x' = f(x)$, avec $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$. Si I est un intervalle de $\mathbb{R}$ et $x : I \to \mathcal{U}$ est une I-solution de (E), pour tout sous-intervalle $J \subset I$ non réduit à un point, la restriction $y = x|_J$ est évidemment une J-solution de (E). Toute L-solution z, avec L intervalle vérifiant $J \subset L$ et $z|_J = y$ sera une solution prolongeant y. La fonction x dont nous sommes partis est évidemment un exemple d'un tel prolongement.

[14] Et la méthode se généralise évidemment aux équations autonomes d'ordre $p > 1$: pour l'équation

$$x^{(p)} = f\left(x, x', \ldots, x^{(p-1)}\right)$$

l'espace dans lequel on travaillera sera de dimension np, avec la fonction inconnue

$$X = \left(x, x', \ldots, x^{(p-1)}\right)$$

[15] C'est à dire une application de classe $\mathcal{C}^2$ vérifiant

$$\forall t \in I \quad (z_2(t), z_2'(t)) \in \mathcal{U} \quad \text{et} \quad z_2''(t) = g(z_2(t), z_2'(t))$$

DÉFINITION 19-2.5 *Soient I_1 et I_2 deux intervalles de $\mathbb{R}$. Si les fonctions x_1 et x_2 sont respectivement une I_1-solution et une I_2-solution de (E), on dit que ces deux solutions se recollent si et seulement si*

$$I_1 \cap I_2 \neq \emptyset \quad et \quad x_1|_{I_1 \cap I_2} = x_2|_{I_1 \cap I_2}$$

On peut alors construire une $(I_1 \cup I_2)$-solution x de (E) (qu'on appellera recollement de x_1 et x_2) en posant

$$\forall t \in I_1 \cup I_2 \quad x(t) = \begin{cases} x_1(t) & \text{si} \quad t \in I_1 \\ x_2(t) & \text{si} \quad t \in I_2 \end{cases}$$

Cette fonction est en effet définie sans ambiguïté. Elle est évidemment de classe C^1 sur $I = I_1 \cup I_2$ et solution de $x' = f(x)$ sur cet intervalle si $I_1 \cap I_2$ est infini (faire un dessin). Dans le cas où $I_1 = (\alpha, \beta]$ et $I_2 = [\beta, \gamma)$, avec $x_1(\beta) = x_2(\beta)$, la fonction x est dérivable à gauche en β avec

$$x'_g(\beta) = x'_1(\beta) = f(x_1(\beta)) = f(x(\beta))$$

et on a un résultat identique à droite. La fonction x est donc bien I-solution de (E).

DÉFINITION 19-2.6 *Une I-solution x de (E) est dite solution maximale si x n'est restriction d'aucune solution de (E) autre qu'elle même.*

Il s'agit donc d'une solution qu'on ne peut prolonger en une solution sur un intervalle contenant strictement I : si $I \subset J$ et y est une J-solution vérifiant $y|_I = x$, on a nécessairement $I = J$.

EXEMPLE 19-2.7 Nous avons vu (conséquence du théorème de Cauchy) que pour une équation linéaire[16]

$$x'(t) = a(t)\,x(t) + b(t)$$

avec $a \in C^0(I, \mathcal{L}(\mathbb{R}^p))$ et $b \in C^0(I, \mathbb{R}^p)$, toute solution définie sur un intervalle $J \subset I$ se prolongeait en une unique solution maximale définie sur I. Par contre, pour l'équation scalaire non linéaire

$$x' = x^2$$

on vérifie facilement que

$$x :]-\infty, 1[\to \mathbb{R} \quad x(t) = -\frac{1}{t-1}$$

est solution maximale, puisque cette solution ne peut évidemment se prolonger à un intervalle contenant 1. On voit donc que la situation est plus complexe dans le cas des équations non linéaires : il y a ici explosion en $t = 1$ pour une solution vérifiant la condition initiale $x(0) = 1$.

[16]Non autonome, il est vrai ... Mais nous verrons (*cf.* section **19-2.2**) que les définitions que nous donnons ici se généralisent aux équations non-autonomes. Si on veut, on peut ici ne considérer que les équations à coefficients constants

$$x' = ax$$

dont les solutions maximales sont définies sur $\mathbb{R}$.

REMARQUE 19-2.8 Pour une équation autonome, la proposition **19-2**.2 montre évidemment que, si $t \mapsto x(t)$ est une I-solution maximale de $x' = f(x)$, l'application $t \mapsto x(t - t_0)$ est également une solution maximale, définie sur $t_0 + I$.

19-2.1.3 Cas des équations scalaires

Commençons par étudier le cas de la dimension 1. On dispose dans ce cas particulier d'outils efficaces pour décrire les solutions : l'utilisation de la monotonie et l'existence de primitives pour une fonction continue sur un intervalle. Considérons donc une équation de la forme

$$x' = f(x)$$

où f est une application de classe C^1 définie sur un intervalle ouvert U à valeurs réelles. Supposons d'abord que nous disposons de deux solutions x_1 et x_2 définies sur deux intervalles I_1 et I_2 contenant un point t_0 tel que

$$x_1(t_0) = x_2(t_0) = x_0 \in U$$

et montrons que x_1 et x_2 se raccordent. Par une translation sur le temps, on peut toujours supposer que $t_0 = 0$. Les solutions se raccordent évidemment si $I_1 = (\alpha, 0]$ et $I_2 = [0, \beta)$. On suppose donc qu'il existe $t > 0$ où les solutions x_1 et x_2 sont définies et montrons que x_1 et x_2 coïncident sur $[0, t]$ (l'étude serait analogue à gauche de 0). Comme x_1 et x_2 sont continues sur le segment $[0, t]$, il existe un segment $K \subset U$ tel que

$$\forall s \in [0, t] \quad x_1(s) \text{ et } x_2(s) \in K$$

Comme f est supposée de classe C^1, la dérivée f' est bornée sur K, et on peut trouver une constante $M > 0$ tellle que $|f'| \leqslant M$ sur K. Par inégalité des accroissements finis, la fonction $y = x_2 - x_1$ vérifie

$$\begin{aligned} y(0) = 0 \quad \text{et} \quad & \forall s \in [0, t] \\ & |y'(s)| = |x_2'(s) - x_1'(s)| = |f(x_2(s)) - f(x_1(s))| \leqslant M\,|y(s)| \end{aligned}$$

On a alors

$$\forall u \in [0, t] \quad |y(u)| = |y(u) - y(0)| \leqslant \int_0^u |y'(s)|\, ds \leqslant M \int_0^u |y(s)|\, ds$$

et on vérifie (par récurrence sur n) comme à la fin de la démonstration du théorème **19-1**.5 que

$$\forall u \in [0, t] \quad |y(u)| \leqslant \frac{(Mu)^n}{n!} \sup_{s \in [0,t]} |y(s)|$$

ce qui montre, en faisant tendre n vers $+\infty$, que y est identiquement nulle sur $[0, t]$: les solutions x_1 et x_2 se raccordent.

Considérons maintenant *a priori* une solution de l'équation définie sur un intervalle I contenant 0, et vérifiant $x(0) = x_0 \in U$.

- Si $f(x_0) = 0$, cette solution se raccorde avec la $\mathbf{R}$-solution évidente constante égale à x_0 (le point x_0 est un équilibre) et donc

$$\forall t \in I \quad x(t) = x_0$$

- Si $f(x_0) \neq 0$, la fonction $t \mapsto f(x(t))$ ne peut s'annuler sur I (puisque s'il existait $t_0 \in I$ avec $f(x(t_0)) = 0$, x serait constante égale à $x(t_0)$ sur I, ce qui est contradictoire avec $x(0) = x_0$).

  On peut supposer $f(x_0) > 0$ (le cas $f(x_0) < 0$ s'étudierait de manière analogue). On a, par le théorème des valeurs intermédiaires,

$$\forall t \in I \quad x'(t) = f(x(t)) > 0$$

  x est strictement croissante sur I et vérifie

$$(*) \quad \forall t \in I \quad \frac{x'(t)}{f(x(t))} = 1$$

  Soit V le sous-intervalle de U contenant x_0

$$V =]x_1, x_2[$$

  avec $x_1 = \sup\{u \in U \mid u < x_0 \text{ et } f(u) = 0\}$ si cet ensemble est non vide, $x_1 = \inf U$ sinon et de même $x_2 = \inf\{u \in U \mid u > x_0 \text{ et } f(u) = 0\}$ ou $x_2 = \sup U$. V est clairement le plus grand sous-intervalle de U contenant x_0 sur lequel f ne s'annule pas. Le théorème des valeurs intermédiaires nous montre que

$$\forall t \in I \quad x(t) \in V$$

  Compte tenu de $(*)$, il est naturel de considérer la primitive F de $\frac{1}{f}$ sur l'intervalle V qui vérifie $F(x_0) = 0$

$$\forall y \in V \quad F(y) = \int_{x_0}^{y} \frac{du}{f(u)}$$

  F est un C^1-difféomorphisme entre V et l'intervalle ouvert $I_0 = F(V)$. On a, en posant $G(t) = F(x(t))$,

$$\forall t \in I \quad G'(t) = 1 \Rightarrow \forall t \in I \quad F(x(t)) - F(x(0)) = F(x(t)) = t$$

  ce qui donne

$$I \subset I_0 \text{ et } \forall t \in I \quad x(t) = F^{-1}(t)$$

  On vérifie enfin que l'application

$$y : I_0 \to V \text{ définie par } y(t) = F^{-1}(t)$$

  est I_0-solution de l'équation, qui vérifie $y(0) = x_0$. Nous avons donc montré que toute solution qui vérifie $x(0) = x_0$ est restriction de y, qui est donc solution maximale. En résumé :

PROPOSITION 19-2.9 L'équation $x' = f(x)$ avec f de classe C^1 sur un intervalle ouvert U de $\mathbb{R}$ possède, pour tout $x_0 \in U$, une unique solution maximale, définie sur un intervalle ouvert contenant 0 et vérifiant la condition initiale $x(0) = x_0$. Toute solution vérifiant cette condition initiale est restriction de cette solution maximale.

Il est intéressant de comprendre pourquoi la solution y définie précédemment est maximale, en regardant pourquoi elle ne peut être prolongée sur un intervalle plus grand que I_0. En nous plaçant ici encore dans le cas où $f(x_0) > 0$, nous avons

$$I_0 =]t_1, t_2[\quad \text{avec} \quad t_1 = \lim_{\substack{x \to x_1 \\ x > x_1}} F(x) \text{ et } t_2 = \lim_{\substack{x \to x_2 \\ x < x_2}} F(x)$$

ce qui donne, puisque $y(t) = F^{-1}(t)$,

$$x_1 = \lim_{\substack{t \to t_1 \\ t > t_1}} y(t) \text{ et } x_2 = \lim_{\substack{t \to t_2 \\ t < t_2}} y(t)$$

Deux cas sont alors envisageables pour le comportement de y pour $t \to t_2$:

- Soit $x_2 = \sup U$ et donc y ne peut se prolonger au delà de t_2 (si t_2 est fini) qu'en sortant de U. On peut aussi avoir $t_2 = +\infty$.

- Soit $x_2 \in U$ est le premier équilibre strictement supérieur à x_0. On a donc $f(x_2) = 0$. Pour $x \in [x_0, x_2[$, nous avons

  $$F(x) = \int_{x_0}^{x} \frac{du}{f(u)}$$

  avec $f > 0$ et

  $$\frac{1}{x_2 - u} \underset{u \to x_2^-}{=} O\left(\frac{1}{f(u)}\right)$$

  puisque f est dérivable en x_2. On en déduit la non-intégrabilité de $\dfrac{1}{f}$ sur $[x_0, x_2[$ et donc

  $$t_2 = \lim_{x \to x_2^-} F(x) = +\infty$$

  On comprend bien pourquoi y est une solution maximale : elle tend vers la position d'équilibre x_2, mais ne l'atteint qu'au bout d'un temps infini.

Sur la figure **19.10**, on observe ces comportements pour trois courbes intégrales de l'équation $x' = 1 - x^2$ pour laquelle les équilibres sont $x_1 = -1$ et $x_2 = 1$. Pour une condition initiale $x_0 \in]-1, 1[$, la solution vérifiant $x(0) = x_0$ est strictement croissante, et tend vers ± 1 en $\pm\infty$. Pour $x_0 < -1$, la solution est décroissante, tend vers -1 en $-\infty$ et explose en un temps fini vers $-\infty$. On peut obtenir le comportement pour $x_0 > 1$ en renversant le cours du temps et en considérant $t \mapsto -x(-t)$.

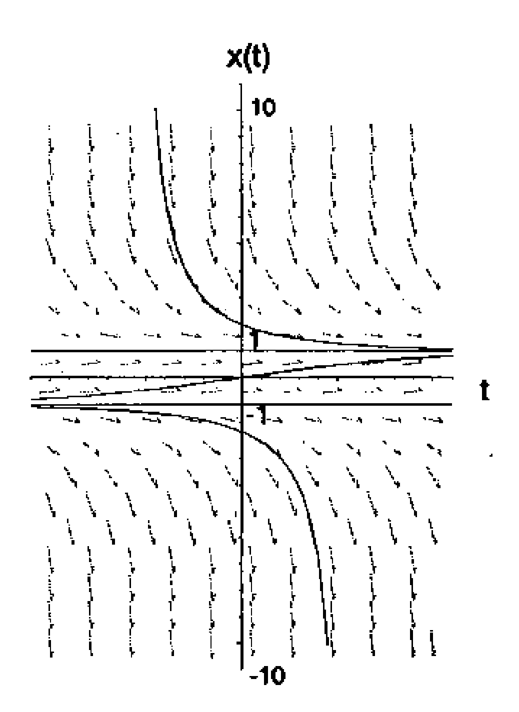


Figure **19.10** – *Champ de vecteurs et trajectoires pour* $x' = 1 - x^2$

REMARQUE 19-2.10 Les raisonnements qui précèdent sont la justification précise du calcul formel, basé sur la séparation des variables

$$\frac{dx}{dt} = f(x) \Leftrightarrow \frac{dx}{f(x)} = dt \Leftrightarrow \int \frac{dx}{f(x)} = t - t_0$$

écriture à proscrire car par trop imprécise !

REMARQUE 19-2.11 Dans le raisonnement qui précède, on a utilisé la caractère lipschitzien (local) de f pour montrer que toute solution de $x' = f(x)$ était restriction d'une unique solution maximale. Si on suppose que f est seulement continue sur l'intervalle ouvert U, on peut montrer que, pour une condition initiale quelconque $x(0) = x_0 \in U$, il existe un intervalle ouvert contenant 0 sur lequel l'équation possède une solution vérifiant cette condition. Mais il n'y a plus en général unicité de cette solution. De même, si toute solution peut se prolonger en une solution maximale, il n'y a pas toujours unicité de ce prolongement :

C'est le cas, par exemple, pour l'équation différentielle

$$(E) : x' = -\left(x^2\right)^{\frac{1}{3}}$$

La fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = -\left(x^2\right)^{\frac{1}{3}}$$

n'est pas lipschitzienne au voisinage de 0. On ne peut donc pas appliquer le théorème précédent. On peut même dire que, pour une condition initiale donnée $x(0) = x_0 > 0$, on trouvera plusieurs solutions maximales :

Soit I un intervalle de $\mathbb{R}$. Remarquons tout d'abord qu'une I-solution x de (E) est décroissante (car $x'(t) \leqslant 0$). S'il existe deux valeurs $t_1 < t_2$ dans I avec $x(t_1) = x(t_2) = 0$, nous aurons x identiquement nulle sur $[t_1, t_2]$. Si nous cherchons des solutions qui ne s'annulent pas sur un intervalle J, nous pourrons appliquer les méthodes que nous venons de voir, puisque f est de classe C^1 sur $U =]0, +\infty[$ ou $]-\infty, 0[$. En particulier, si x est une I-solution avec 0 intérieur à I et $x(0) = x_0 > 0$, nous aurons par décroissance de x

$$\forall t \in I \cap \mathbb{R}^- \quad x(t) > 0$$

et donc

$$t < 0 \Rightarrow \int_t^0 \frac{x'(s)}{x^{\frac{2}{3}}(s)} ds = -t \Rightarrow 3\left[x^{\frac{1}{3}}(t) - x_0^{\frac{1}{3}}\right] = -t \Rightarrow x(t) = \frac{(t_0 - t)^3}{27}$$

où $t_0 = 3x_0^{\frac{1}{3}} > 0$. Cette expression donne en fait une solution x de (E) dans l'intervalle $J =]-\infty, t_0[$, strictement positive et vérifiant $x(0) = x_0$. La proposition **19-2**.9 montre (puisque f est de classe C^1 sur $U =]0, +\infty[$) que toute solution y de (E) **ne s'annulant pas** et vérifiant $y(0) = x_0$ est restriction[17] de x. Il en résulte facilement qu'une solution y de (E) vérifiant la même condition initiale se recolle[18] avec x.

Que se passe-t-il si nous voulons prolonger x au-delà de t_0 ? Nous pouvons raccorder x avec la solution identiquement nulle sur $[t_0, +\infty[$. Ou, si $t_1 > t_0$, prolonger par

$$\forall t \in [t_0, t_1] \quad x(t) = 0 \quad \text{et} \quad \forall t > t_1 \quad x(t) = \frac{(t_1 - t)^3}{27}$$

ou encore prendre

$$\forall t \in \mathbb{R} \quad x(t) = \frac{(t_0 - t)^3}{27}$$

Nous obtenons ainsi plusieurs solutions maximales (en fait définies sur $\mathbb{R}$) vérifiant $x(0) = x_0$. De plus, si $t \mapsto y(t)$ est I-solution de (E), il est clair que l'application $t \mapsto y(-t)$ est $(-I)$-solution. On peut ainsi voir que toutes les solutions maximales vérifiant la condition $x(0) = x_0$ sont d'une des formes décrites plus haut (exercice).

[17] Puisque x est clairement solution maximale de

$$x' = f(x) \quad \text{avec} \quad x > 0$$

[18] Parce que si y est K-solution vérifiant $y(0) = x_0$, y ne peut s'annuler sur $K \cap]-\infty, t_0[$. On arriverait en effet à une contradiction en considérant le plus petit zéro > 0 de y.

19-2.1.4 Existence et unicité locale d'une solution

Nous étudions à présent le cas général, avec $f \in C^1(\mathcal{U}, \mathbb{R}^p)$ et $\mathcal{U}$ ouvert de $\mathbb{R}^p$. Nous allons, pour approcher des solutions de

$$(E) : x' = f(x)$$

utiliser le **schéma d'Euler**, basé sur la remarque suivante : si $x_0 \in \mathcal{U}$ et si I est un intervalle ouvert contenant 0, une I-solution de (E) vérifiant $x(0) = x_0$ vérifiera, pour t "petit"

$$x(t) \approx x(0) + tx'(0) = x_0 + tf(x_0)$$

Si $h > 0$ est choisi assez petit, on approchera donc la solution sur $[0, h]$ par cette expression, en aboutissant en

$$x_1 = x_0 + hf(x_0) \approx x(h)$$

On poursuit ensuite de la même façon sur $[h, 2h]$ par

$$x(h+t) \approx x(h) + tf(x(h)) \approx x_1 + tf(x_1)$$

ce qui amène à

$$x(2h) \approx x_2 = x_1 + hf(x_1)$$

On construit donc de proche en proche une suite de points $(x_k)_{k=0,1,2,\cdots}$ avec

$$(*) \quad x_{k+1} = x_k + hf(x_k)$$

permettant d'envisager une ligne polygonale continue dans $\mathcal{U}$, représentant une fonction affine par morceaux y_h vérifiant

$$\forall t \in [kh, (k+1)h] \quad y_h(t) = x_k + \frac{t - kh}{h}(x_{k+1} - x_k)$$

On opérera bien sûr de la même façon à gauche de 0, en posant

$$x_{-1} = x_0 - hf(x_0) \quad \text{et} \quad x_k = x_{k+1} - hf(x_{k+1}) \text{ pour } k = -2, -3, \ldots$$

et en définissant y_h par interpolation. Il y a évidemment une difficulté pour itérer suffisamment le procédé : pour construire x_{k+1} à partir de x_k à l'aide de la formule $(*)$, il faut disposer de $f(x_k)$ et donc savoir que $x_k \in \mathcal{U}$. Ceci va nous amener, dans un premier temps, à ne pas être trop ambitieux sur la longueur de l'intervalle contenant 0 où nous chercherons une solution approchée de (E) :

LEMME 19-2.12 Soit $f : \mathcal{U} \to \mathbb{R}^p$ de classe C^1 et $x_0 \in \mathcal{U}$. Il existe un $T > 0$ tel que, pour tout $n \in \mathbb{N}^*$, la suite (x_k) du schéma d'Euler associé au pas $h = \frac{T}{n}$ soit définie pour tout k compris entre $-n$ et n. Cette suite permet alors de définir une fonction continue affine par morceaux

$$z_n = y_{\frac{T}{n}} : [-T, T] \to \mathcal{U}$$

"solution approchée" de l'équation

$$x' = f(x)$$

pour la condition initiale $x(0) = x_0$.

Démonstration : $\mathbb{R}^p$ est muni d'une de ses normes usuelles. Comme $x_0 \in \mathcal{U}$ ouvert, on commence par déterminer un $\alpha > 0$ avec

$$B(x_0, \alpha] \subset \mathcal{U}$$

La fonction f est continue sur le compact $B(x_0, \alpha]$, et donc

$$\exists M > 0 \quad \forall x \in B(x_0, \alpha] \quad \|f(x)\| \leqslant M$$

Montrons que

$$T = \frac{\alpha}{M}$$

répond à la question : prenons $n \geqslant 1$, $h = \dfrac{T}{n}$ et montrons que le schéma d'Euler correspondant peut être itéré jusqu'à $k = n$, avec $x_k \in B(x_0, \alpha]$ pour $k = 0, \dots, n$ (l'étude serait analogue pour les valeurs négatives de k). Si ce n'est pas le cas, il existe k_0 entier vérifiant $0 \leqslant k_0 \leqslant n - 1$ avec

$$x_0, x_1, \dots, x_{k_0} \in B(x_0, \alpha] \text{ et } \|x_{k_0+1} - x_0\| > \alpha$$

Comme pour $k \leqslant k_0$

$$x_{k+1} = x_k + hf(x_k), \text{ nous avons } \|x_{k+1} - x_k\| = \frac{T}{n}\|f(x_k)\| \leqslant \frac{MT}{n}$$

puisque $x_k \in B(x_0, \alpha]$. Par inégalité triangulaire, nous avons

$$\|x_{k_0+1} - x_0\| \leqslant \sum_{k=0}^{k_0} \|x_{k+1} - x_k\| \leqslant (k_0 + 1)\frac{MT}{n} \leqslant MT = \alpha$$

ce qui contredit l'hypothèse faite sur k_0. Il reste à voir que, puisque la boule $B(x_0, \alpha]$ est convexe, si x_k et $x_{k+1} \in B(x_0, \alpha]$, le segment $[x_k, x_{k+1}]$ est inclus dans $B(x_0, \alpha]$, et la fonction affine par morceaux

$$z_n = y_{\frac{T}{n}} : [-T, T] \to \mathbb{R}^n$$

dont l'image est la réunion des segments $[x_k, x_{k+1}]$, $k = -n, \dots, n-1$ prend ses valeurs dans $B(x_0, \alpha]$. ■

Le chemin est à présent tout tracé pour construire une solution de (E) sur $[-T, T]$ vérifiant la condition initiale $x(0) = x_0$: nous allons montrer que la suite de fonctions $(z_n)_{n \geqslant 1}$ converge uniformément sur $[-T, T]$ vers une fonction z continue. Il restera à voir que z est $\mathcal{C}^1$ et vérifie (E). La convergence uniforme se prouve à l'aide du lemme suivant :

LEMME 19-2.13 Si $(z_n)_{n \geqslant 1}$ est la suite de solutions approchées définies dans le lemme précédent, il existe une constante $B > 0$ telle que

$$\forall n \geqslant 1 \quad \forall t \in [-T, T] \quad \|z_n'(t) - f(z_n(t))\| \leqslant \frac{B}{n}$$

Démonstration : La fonction z_n est continue et C^1 par morceaux. Dans l'inégalité qui précède, en un point t de non dérivabilité, $z_n'(t)$ représente une dérivée à droite ou à gauche. Plaçons-nous sur le segment $[kh,(k+1)h] = \left[k\frac{T}{n},(k+1)\frac{T}{n}\right]$. On a alors, avec les notations du début de cette section

$$z_n(t) = x_k + (t-kh)f(x_k)$$

donc, par construction

$$z_n'(t) = f(x_k) \quad \text{et} \quad \|z_n(t) - x_k\| \leqslant h\|f(x_k)\| \leqslant \frac{T}{n}M = \frac{\alpha}{n}$$

où α et M ont la même signification que dans la démonstration du lemme précédent. Comme f est de classe C^1, on peut trouver une constante $K > 0$ telle que

$$\forall x \in B(x_0,\alpha] \quad \|df_x\| \leqslant K$$

(continuité de l'application $x \mapsto df_x$ sur le compact $B(x_0,\alpha]$, la norme utilisée est ici la norme des endomorphismes de $\mathbb{R}^p$ subordonnée à la norme choisie sur $\mathbb{R}^p$). On a alors, puisque $z_n(t)$ et $x_k \in B(x_0,\alpha]$, par inégalité des accroissements finis,

$$\|z_n'(t) - f(z_n(t))\| = \|f(x_k) - f(z_n(t))\| \leqslant K\|z_n(t) - x_k\| \leqslant \frac{K\alpha}{n}$$

et on peut donc choisir $B = K\alpha$. ∎

Cette inégalité montre que, pour n suffisamment grand, z_n n'est pas loin de vérifier (E). Considérons à présent deux entiers $m > n \geqslant 1$. On a par inégalité triangulaire

$$\forall t \in [-T,T] \quad \|z_m'(t) - z_n'(t)\| \leqslant$$
$$\|f(z_m(t)) - f(z_n(t))\| + \frac{B}{m} + \frac{B}{n} \leqslant K\|z_m(t) - z_n(t)\| + \frac{2B}{n}$$

où, comme précédemment, K majore la norme de la différentielle de f sur la boule $B(x_0,\alpha]$. La fonction (continue et C^1 par morceaux) $u = z_m - z_n$ vérifie donc

$$\forall t \in [-T,T] \quad \|u'(t)\| \leqslant K\|u(t)\| + \frac{2B}{n}$$

et nous pouvons lui appliquer le lemme suivant (dit lemme de Gronwall) :

LEMME 19-2.14 Soit I un intervalle de $\mathbb{R}$ et $v : I \to \mathbb{R}^p$ continue et C^1 par morceaux telle qu'il existe deux constantes $a > 0$ et $b \geqslant 0$ avec

$$\forall t \in I \quad \|v'(t)\| \leqslant a\|v(t)\| + b$$

alors, pour $t_0 \in I$ quelconque

$$\forall t \in I \quad \|v(t)\| \leqslant \|v(t_0)\| e^{a|t-t_0|} + \frac{b}{a}\left(e^{a|t-t_0|} - 1\right)$$

Démonstration : Nous considérerons le cas où $t \geqslant t_0$, le cas symétrique se traitant de manière analogue. On a

$$\forall t \in I \quad t \geqslant t_0 \Rightarrow \|v(t)\| \leqslant \|v(t_0)\| + \int_{t_0}^{t} \|v'(u)\|\, du$$
$$\leqslant \|v(t_0)\| + a\int_{t_0}^{t} \|v(u)\|\, du + b(t - t_0)$$

La fonction de classe C^1 définie par

$$F(t) = a\int_{t_0}^{t} \|v(u)\|\, du + b(t - t_0) + \|v(t_0)\|$$

vérifie donc

$$\forall t \geqslant t_0 \quad F'(t) - aF(t) \leqslant b$$

Ceci montre que la fonction $t \mapsto e^{-at}\left[F(t) + \dfrac{b}{a}\right]$ est décroissante pour $t \geqslant t_0$, et donc

$$\forall t \geqslant t_0 \quad F(t) \leqslant e^{a(t-t_0)}\left[F(t_0) + \frac{b}{a}\right] - \frac{b}{a}$$
$$= \|v(t_0)\|\, e^{a(t-t_0)} + \frac{b}{a}\left(e^{a(t-t_0)} - 1\right)$$

On en déduit le résultat, puisque $\|v(t)\| \leqslant F(t)$. ■

En appliquant ce lemme à la fonction u, avec $t_0 = 0$ et $u(0) = 0$ (puisque z_n et z_m vérifient la même condition initiale), on obtient

$$\forall t \in [-T,T] \quad m \geqslant n \Rightarrow \|z_m(t) - z_n(t)\| \leqslant \frac{2B}{nK}\left(e^{K|t|} - 1\right) \leqslant \frac{2B}{nK}\left(e^{KT} - 1\right)$$

Cette majoration montre que la suite de fonctions $(z_n)_{n\geqslant 1}$ est uniformément de Cauchy sur $[-T,T]$, et converge donc uniformément sur ce segment vers une fonction continue

$$z = \lim_{n\to+\infty} z_n : [-T,T] \to B(x_0, \alpha]$$

De plus, pour $t \in [-T,T]$, on a, puisque z_n est continue et C^1 par morceaux

$$(**) \quad \forall n \in \mathbb{N}^* \quad z_n(t) = x_0 + \int_0^t z_n'(s)\, ds$$

Les majorations

$$\forall s \in [-T,T] \quad \|f(z_n(s)) - f(z(s))\| \leqslant K\|z_n(s) - z(s)\|$$
$$\text{et} \quad \|z_n'(s) - f(z_n(s))\| \leqslant \frac{B}{n}$$

montrent la convergence uniforme sur $[-T,T]$ de la suite de fonctions $(z_n')_{n\geqslant 1}$ vers $f\circ z$. Il en résulte qu'on peut passer à la limite dans (**) pour obtenir

$$\forall t\in[-T,T]\quad z(t)=x_0+\int_0^t f(z(s))\,ds$$

La fonction z est donc de classe $\mathcal{C}^1$ sur $[-T,T]$ et est solution de (E) sur ce segment, en vérifiant la condition initiale $z(0)=x_0$.

Nous avons ainsi obtenu un résultat d'existence locale d'une solution de (E) pour une condition initiale donnée. En fait, les lemmes démontrés précédemment permettent d'obtenir également un résultat d'unicité (locale) : considérons en effet un intervalle I contenant 0 et $x:I\to\mathcal{U}$ une I-solution de (E) vérifiant $x(0)=x_0$ et montrons que x coïncide avec z sur l'intervalle $J=I\cap[-T,T]$. On pourra alors recoller x et z :

- On voit d'abord que, pour tout $t\in J$, nous avons $x(t)\in B(x_0,\alpha]$. Si ce n'était pas le cas, il existerait un $t_1>0$ dans J avec $\|x(t_1)-x_0\|>\alpha$ (l'étude avec un $t_1<0$ serait analogue). Considérons alors

$$t_0=\inf\{t>0\mid t\in J \text{ et } \|x(t)-x_0\|\geqslant\alpha\}$$

Pour des raisons de continuité, nous avons $0<t_0<t_1\leqslant T$ et

$$\forall t\in[0,t_0]\quad x(t)\in B(x_0,\alpha]$$

avec $\|x(t_0)-x_0\|=\alpha$. Mais on a alors

$$\|x(t_0)-x_0\|=\left\|\int_0^{t_0} f(x(s))\,ds\right\|\leqslant\int_0^{t_0}\|f(x(s))\|\,ds\leqslant Mt_0<MT=\alpha$$

ce qui est contradictoire avec la définition de t_0. On a donc bien

$$\forall t\in J\quad x(t)\in B(x_0,\alpha]$$

- Nous avons alors

$$\forall n\geqslant 1\ \forall t\in J\ \ \|x'(t)-z'(t)\|=\|f(x(t))-f(z(t))\|\leqslant K\,\|x(t)-z(t)\|$$

et le lemme de Gronwall donne alors

$$\forall t\in J\quad \|x(t)-z(t)\|\leqslant\|x(0)-z(0)\|\,e^{K(t-t_0)}$$

ce qui donne $x(t)=z(t)$, puisque $x(0)-z(0)=0$.

Résumons l'étude qui précède par l'énoncé de la version locale du théorème de Cauchy-Lipschitz :

THÉORÈME 19-2.15 Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^p$, $f:\mathcal{U}\to\mathbb{R}^p$ une application de classe $\mathcal{C}^1$ et x_0 un point de $\mathcal{U}$. Il existe un réel $T>0$ tel que

- **Il existe une solution z de (E) : $x'=f(x)$ sur $[-T,T]$ vérifiant la condition initiale $z(0)=x_0$.**
- **Toute solution y de (E) sur un intervalle I contenant 0 et vérifiant $y(0)=x_0$ se raccorde avec z, c'est à dire coïncide avec z sur $[-T,T]\cap I$.**

EXERCICE 19-2.16 Pour l'équation linéaire à coefficients constants

$$X' = AX \quad \text{avec} \quad A \in \mathcal{M}_p(\mathbb{R})$$

montrer (en utilisant la démonstration précédente) que le schéma d'Euler converge sur $[0,1]$ et en déduire que

$$\exp(A) = \lim_{n \to +\infty} \left(I_p + \frac{A}{n} \right)^n$$

REMARQUE 19-2.17 Dans la démonstration du théorème **19-2**.15, nous n'avons utilisé le caractère C^1 de f que d'une manière affaiblie : on peut en fait supposer que f est simplement **localement lipschitzienne sur l'ouvert** $\mathcal{U}$, ce qui signifie :

Pour tout compact $\mathcal{K} \subset \mathcal{U}$, la restriction de f à $\mathcal{K}$ est lipschitzienne (avec une constante de Lipschitz qui dépend *a priori* du compact $\mathcal{K}$ choisi).

19-2.1.5 Solutions maximales : théorème de Cauchy-Lipschitz

Le théorème **19-2**.15 donne immédiatement :

THÉORÈME 19-2.18 Soient $f : \mathcal{U} \to \mathbb{R}^p$ est de classe C^1 et I_1, I_2 deux intervalles de $\mathbb{R}$ avec $I_1 \cap I_2 \neq \emptyset$. Si $x_1 : I_1 \to \mathcal{U}$ et $x_2 : I_2 \to \mathcal{U}$ sont deux solutions de

$$(E) : x' = f(x)$$

qui coïncident en un point de $I_1 \cap I_2$, alors x_1 et x_2 se raccordent.

Démonstration : L'ensemble $F = \{t \in I_1 \cap I_2 \mid x_1(t) = x_2(t)\}$ est non vide, et c'est évidemment un fermé de $I_1 \cap I_2$. Si $t_0 \in F$, et si $x_0 = x_1(t_0) = x_2(t_0)$, le théorème **19-2**.15 (après une translation sur le temps) assure l'existence d'un $T > 0$ et d'une solution

$$z : [t_0 - T, t_0 + T] \to \mathcal{U}$$

telle que z coïncide avec x_1 sur $[t_0 - T, t_0 + T] \cap I_1$ et avec x_2 sur $[t_0 - T, t_0 + T] \cap I_2$. On a donc

$$F \supset [t_0 - T, t_0 + T] \cap (I_1 \cap I_2)$$

et F est voisinage (relatif) de t_0 dans $I_1 \cap I_2$. F est donc à la fois ouvert et fermé non vide du connexe $I_1 \cap I_2$. Donc $F = I_1 \cap I_2$: x_1 et x_2 se raccordent. ■

Ce résultat de raccordement permet d'obtenir le théorème (dit de Cauchy-Lipschitz) d'existence et unicité d'une solution maximale vérifiant une condition initiale donnée :

THÉORÈME 19-2.19 Soit $f : \mathcal{U} \to \mathbb{R}^p$ de classe C^1 (ou simplement localement lipschitzienne sur $\mathcal{U}$). Toute solution de $(E) : x' = f(x)$ se prolonge de manière unique en une solution maximale, définie sur un intervalle ouvert. Pour tout $x_0 \in \mathcal{U}$ et $t_0 \in \mathbb{R}$, l'équation (E) possède une unique solution maximale vérifiant la condition initiale

$$x(t_0) = x_0$$

Démonstration : Soit I_0 un intervalle de $\mathbb{R}$ et y une I_0-solution de (E). Considérons l'ensemble X formé des intervalles J de $\mathbb{R}$ tels que

$$I_0 \subset J \text{ et } y \text{ se prolonge en une } J\text{-solution } y_J \text{ de } (E)$$

(ensemble non vide, puisqu'il contient I_0). Remarquons d'abord que, pour $J \in X$, le prolongement y_J de y est unique (théorème **19-2**.18). Posons

$$I = \bigcup_{J \in X} J$$

I est réunion d'intervalles contenant I_0, et est donc un intervalle. Montrons qu'il est ouvert : s'il s'écrit par exemple $I = (\alpha, \beta]$ avec $\alpha \in \mathbb{R} \cup \{-\infty\}$ et $\beta \in \mathbb{R}$, il existe $J \in X$ avec $\beta \in J$. L'intervalle J est donc de la forme $(\gamma, \beta]$, il contient I_0 et y s'y prolonge en une solution y_J. Nous allons voir qu'on peut prolonger y_J en une solution à droite de β, ce qui est évidemment contradictoire avec $\beta = \sup I$. Pour ce faire, il suffit de considérer $y_0 = y_J(\beta) \in \mathcal{U}$, et d'utiliser le théorème **19-2**.15 qui nous assure l'existence d'un $T > 0$ et d'une solution $z : [\beta - T, \beta + T] \to \mathcal{U}$ vérifiant $z(\beta) = y_0 = y_J(\beta)$. D'après le théorème **19-2**.18, les solutions y_J et z se raccordent, ce qui permet de prolonger y_J au delà de β.

L'intervalle I est donc ouvert. Remarquons qu'en adaptant le raisonnement précédent, nous voyons que tout point $t \in I$ appartient à un intervalle **ouvert** sur lequel est définie une solution prolongeant y. Définissons une fonction $x : I \to \mathcal{U}$ par

$$\forall t \in I \quad x(t) = y_J(t) \text{ si } t \in J \text{ élément de } X$$

Cette fonction est parfaitement définie car si t appartient à deux intervalles J_1 et J_2, les solutions y_{J_1} et y_{J_2} se raccordent (puisqu'elles sont égales sur I_0) et donc $y_{J_1}(t) = y_{J_2}(t)$. De plus, x est de classe C^1 et est I-solution de (E), puisque tout $t \in I$ appartient à un intervalle ouvert sur lequel x coïncide avec une solution de (E). Par construction, x est maximale et prolonge y.

Pour terminer, si $(t_0, x_0) \in \mathbb{R} \times \mathcal{U}$, la solution locale de (E) vérifiant $z(t_0) = x_0$ donnée par le théorème **19-2**.15 se prolonge comme on vient de le voir en une solution maximale, forcément unique d'après le théorème **19-2**.18. ∎

REMARQUE 19-2.20 Pour une équation autonome du second ordre

$$(E_2) : x'' = g(x, x')$$

avec g de classe C^1 sur un ouvert $\mathcal{U}$ de $\mathbb{R}^p \times \mathbb{R}^p$, on aura le même résultat pour une condition initiale

$$t_0 \in \mathbb{R} \quad \text{et} \quad (x_0, x_1) \in \mathcal{U}$$

Il existe une unique solution maximale x de (E_2), définie sur un intervalle ouvert et vérifiant

$$\begin{cases} x(t_0) = x_0 \\ x'(t_0) = x_1 \end{cases}$$

Ceci se voit en appliquant le théorème de Cauchy-Lipschitz au système du premier ordre équivalent Σ_{E_2}(voir ce qui suit la définition **19-2.4**). On vérifie facilement que le résultat subsiste si g est simplement localement lipschitzienne, et se généralise de manière évidente aux équations autonomes d'ordre $p > 2$.

REMARQUE 19-2.21 Nous nous sommes occupés, dans ce qui précède, de la résolution du **problème de Cauchy** : recherche de solutions (d'abord locales, puis maximales) de $x' = f(x)$ pour une **condition initiale donnée**. Par exemple, un solution du système (linéaire) d'équations (scalaires)

$$(S) : \begin{cases} x' = y \\ y' = -\omega^2 x \end{cases}$$

(où ω est un scalaire non nul) est caractérisée par la donnée du couple

$$(x(0), y(0)) = (\alpha, \beta) \in \mathbb{R}^2$$

Comme (S) est le système du premier ordre associé à l'équation du second ordre

$$x'' + \omega^2 x = 0$$

cette solution est évidemment

$$x(t) = \alpha \cos \omega t + \frac{\beta}{\omega} \sin \omega t \quad \text{et} \quad y(t) = x'(t)$$

Il ne faut pas croire cependant que deux "conditions scalaires" apparemment "indépendantes" permettent en général de caractériser les solutions de (S) : discuter par exemple l'existence de solutions de (S) vérifiant les "conditions aux limites"

$$x(0) = 0 \quad \text{et} \quad x(1) = \gamma$$

en fonction des paramètres γ et ω.

19-2.1.6 Interprétation géométrique

Le théorème de Cauchy-Lipschitz se traduit géométriquement à l'aide de la notion de courbe intégrale et de trajectoire (définition **19-2**.3). Si $f : \mathcal{U} \to \mathbb{R}^p$ est de classe C^1 et (E) est l'équation $x' = f(x)$, considérons d'abord le cas d'une condition initiale $x_0 \in \mathcal{U}$ correspondant à un équilibre, c'est à dire telle que

$$f(x_0) = 0$$

Nous connaissons alors la $\mathbb{R}$-solution constante

$$\forall t \in \mathbb{R} \quad x(t) = x_0$$

Le théorème de Cauchy-Lipschitz nous montre qu'une solution qui prendrait cette valeur serait forcément constante sur son intervalle de définition. En d'autres termes, le singleton $\{x_0\}$ est une trajectoire solution de (E), et toute autre trajectoire ne peut contenir le point x_0.

De même, considérons deux trajectoires Γ_1 et Γ_2 de solutions de (E) que nous notons $t \mapsto x_1(t)$ et $t \mapsto x_2(t)$ définies sur des intervalles respectifs I_1 et I_2, et supposons que ces deux trajectoires se coupent en un point $x_0 \in \mathcal{U}$:

$$\exists t_1 \in I_1 \quad \exists t_2 \in I_2 \quad x_1(t_1) = x_2(t_2) = x_0$$

La solution de (E) définie sur $(t_1 - t_2) + I_2$ par

$$z(t) = x_2(t + t_2 - t_1)$$

est alors définie en t_1 et vérifie

$$z(t_1) = x_1(t_1)$$

Les solutions z et x_1 se recollent, et on peut donc considérer que Γ_1 et Γ_2 sont deux trajectoires incluses dans la même trajectoire "maximale", celle correspondant à l'unique solution maximale prolongeant x_1 et z. (La solution x_2 en est une restriction, après translation sur le temps).

PROPOSITION 19-2.22 Deux trajectoires maximales de solutions de (E) sont disjointes ou confondues. Les différentes trajectoires maximales réalisent donc une partition de l'ouvert $\mathcal{U}$.

EXERCICE 19-2.23 Si x est I-solution de $x' = f(x)$ avec $f \in C^1(\mathcal{U}, \mathbb{R}^p)$ et s'il existe $t_1 \neq t_2 \in I$ avec $x(t_1) = x(t_2)$, montrer que la solution maximale prolongeant x est définie sur $\mathbb{R}$ et que cette solution est périodique de période $|t_2 - t_1|$. (En d'autres termes, si une trajectoire se "recoupe", c'est une trajectoire d'une solution périodique). Dans le cas $p = 1$, montrer que les seules solutions périodiques de $x' = f(x)$ sont constantes. Ce n'est évidemment plus vrai en dimension 2, comme le montre l'exemple de la remarque **19-2**.21.

19-2.1.7 Comportement "au bord"

PROPOSITION 19-2.24 Soit $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$ et x une solution de $x' = f(x)$ définie sur un intervalle $I = [a, b[$ avec $b < +\infty$. Si la limite

$$l = \lim_{\substack{t \to b \\ t < b}} x(t)$$

existe et appartient à $\mathcal{U}$, alors x n'est pas maximale.

Démonstration : Il suffit de voir que x peut se prolonger en une solution sur $[a, b]$: on peut prolonger x par continuité en posant $x(b) = l$, et comme

$$\lim_{\substack{t \to b \\ t < b}} x'(t) = \lim_{\substack{t \to b \\ t < b}} f(x(t)) = f(l)$$

le prolongement est C^1 sur $[a, b]$ avec $x'(b) = f(l) = f(x(b))$ (théorème de prolongement d'une dérivée). Ce prolongement est alors solution de (E) sur $[a, b]$. On peut alors bien sûr recoller ensuite avec une solution locale y vérifiant $y(b) = l$ pour prolonger au delà de b. ■

EXERCICE 19-2.25 Déduire de ce qui précède que, si I est majoré et s'il existe un compact $\mathcal{K} \subset \mathcal{U}$ tel que

$$\forall t \in I \quad x(t) \in \mathcal{K}$$

alors x n'est pas maximale.

Dans l'exercice qui suit, nous voyons (comme dans le cas de la dimension 1) ce qui empêche une solution maximale définie sur un intervalle ouvert $]\alpha, \beta[$ de se prolonger au delà de β. C'est essentiellement parce que $\beta = +\infty$ (!), ou parce que la trajectoire "s'échappe à l'infini" (dans l'ouvert $\mathcal{U}$, c'est à dire finit par sortir de tout compact de $\mathcal{U}$ sans jamais y revenir) :

EXERCICE 19-2.26 Soit $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$ et x une solution de $x' = f(x)$ définie sur un intervalle $I = [a, b[$ avec $b < +\infty$. Montrer que, s'il existe une suite $(t_n)_{n \in \mathbb{N}}$ de points de I, strictement croissante et convergeant vers b telle que

$$\lim_{n \to +\infty} x(t_n) = x_0$$

existe et appartienne à $\mathcal{U}$, alors x n'est pas maximale.

Indication : si on pose $x_n = x(t_n)$, montrer qu'on peut trouver un $\alpha > 0$ et un entier n_0 tel que

$$\forall n \geqslant n_0 \quad B\left(x_n, \frac{\alpha}{2}\right] \subset B(x_0, \alpha] \subset \mathcal{U}$$

et, en utilisant les résultats intermédiaires de la démonstration du théorème de Cauchy-Lipschitz, montrer que x peut être prolongée au delà de b.

En déduire que si x est une solution maximale de (E) dont l'intervalle de définition I est borné supérieurement, pour tout $\mathcal{K}$ compact de $\mathcal{U}$

$$\exists t_0 \in I \quad \forall t \geqslant t_0 \quad x(t) \notin \mathcal{K}$$

(en renversant le cours du temps, on aurait évidemment un résultat analogue à la borne inférieure de I).

On voit donc que, si x est une I solution maximale de (E), son graphe

$$\Gamma_x = \{(t, x(t)) \mid t \in I\}$$

n'est inclus dans aucun compact de $\mathbb{R} \times \mathcal{U}$.

EXERCICE 19-2.27 Si $f : \mathbb{R}^p \to \mathbb{R}^p$ est de classe C^1 et vérifie

$$\exists K, L > 0 \quad \forall x \in \mathbb{R}^p \quad \|f(x)\| \leqslant K\|x\| + L$$

montrer que les solutions maximales de $x' = f(x)$ sont définies sur $\mathbf{R}$. Lorsque f est K-lipschitzienne, il en est de même (l'hypothèse f est de classe C^1 est alors superflue, comme le montre la démonstration du théorème de Cauchy-Lipschitz).

Cet exercice montre que, lorsque le champ de vecteurs $x \mapsto f(x)$ est défini sur $\mathbb{R}^p$, on peut parfois prévoir que les solutions maximales seront définies sur $\mathbf{R}$ tout entier.

19-2.2 Cas des équations non-autonomes

On considère maintenant des équations de la forme

$$(E) : x' = g(t, x)$$

où g est une application de classe C^1 d'un ouvert $\mathcal{U} \subset \mathbb{R} \times \mathbb{R}^p$ à valeurs dans $\mathbb{R}^p$. Si I est un intervalle de $\mathbf{R}$ non réduit à un point, on appelle I-solution de (E) toute application $y : I \to \mathbf{R}^p$ dérivable (en fait forcément de classe C^1) telle que

$$\forall t \in I \quad (t, y(t)) \in \mathcal{U} \quad \text{et} \quad y'(t) = g(t, y(t))$$

Les notions de restriction, recollement, solution maximale sont les mêmes que dans le cas des équations autonomes. On peut de fait ramener l'étude de (E) à celle d'une équation autonome, un peu comme lorsqu'on transforme une équation du second ordre en un système du premier ordre : on augmente la dimension de "l'espace des phases", en interprétant la variable t (le temps) comme une variable d'espace :

PROPOSITION 19-2.28 Soit $(t_0, x_0) \in \mathcal{U}$ et I un intervalle (non réduit à un point) contenant t_0. Une fonction $x : I \to \mathbb{R}^p$ est I-solution de (E) vérifiant la condition initiale $x(t_0) = x_0$ si et seulement si l'application

$$X : I \to \mathbb{R}^{p+1} \quad t \mapsto X(t) = (t, x(t))$$

est telle que

$$\forall t \in I \quad X(t) \in \mathcal{U}$$

et est solution de l'équation différentielle

$$(E') : X'(t) = F(X(t))$$

pour la condition initialle

$$X(t_0) = (t_0, x_0)$$

où $F : \mathcal{U} \to \mathbb{R} \times \mathbb{R}^p$ est définie par

$$\forall \tau \in \mathbb{R} \quad \forall y \in \mathbb{R}^p \quad (\tau, y) \in \mathcal{U} \Rightarrow F(\tau, y) = (1, g(\tau, y))$$

Démonstration : C'est évident. On peut noter que, réciproquement, si

$$Z : I \ni t \mapsto Z(t) = (T(t), y(t))$$

est I-solution de (E') avec $Z(t_0) = (t_1, y_1) \in \mathcal{U}$ pour un $t_0 \in I$, nous avons

$$\forall t \in I \quad T'(t) = 1 \text{ et } y'(t) = g(T(t), y(t))$$

ce qui donne

$$\forall t \in I \quad T(t) = t - (t_0 - t_1)$$

La fonction x définie sur $J = I - (t_0 - t_1)$ par $z(t) = y(t + t_0 - t_1)$ vérifie donc

$$\begin{aligned} \forall t \in J \quad z'(t) &= y'(t + t_0 - t_1) \\ &= g(T(t + t_0 - t_1), y(t + t_0 - t_1)) = g(t, z(t)) \end{aligned}$$

C'est une solution de (E) pour la condition initiale $z(t_1) = y_1$. Toute solution de (E') permet donc, après translation sur le temps, de reconstituer une solution de (E). ■

Si $g \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$, on a évidemment $F \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^{p+1})$, et le théorème de Cauchy-Lipschitz appliqué à (E') donne immédiatement :

THÉORÈME 19-2.29 Si $\mathcal{U}$ est un ouvert de $\mathbb{R}^{p+1}$ et $g \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^p)$, pour tout couple (t_0, x_0) dans $\mathcal{U}$, l'équation

$$(E) : x' = g(t, x)$$

possède une unique solution maximale x, définie sur un intervalle ouvert, vérifiant la condition initiale $x(t_0) = x_0$. Toute solution de (E) qui vérifie cette condition initiale est restriction de cette solution maximale.

Suivre dans l'espace des phases $\mathbb{R}^{p+1}$ la trajectoire de $t \mapsto (t, x(t))$, c'est en fait considérer la "représentation graphique" de $t \mapsto x(t)$. Le théorème qui précède nous dit donc simplement que les représentations graphiques des solutions maximales de (E) forment une partition de l'ouvert $\mathcal{U}$.

L'exercice **19-2**.26 appliqué à l'équation (E') montre que, si x est solution maximale de (E) définie sur $I =]\alpha, \beta[$ et $\mathcal{K}$ est un compact inclus dans $\mathcal{U}$, on peut trouver $t_1 \in \mathcal{U}$ tel que

$$\forall t \in I \quad t > t_1 \Rightarrow (t, x(t)) \notin \mathcal{K}$$

REMARQUE 19-2.30 Un examen attentif des démonstrations de la section **19-2.1.4** nous montrerait que le théorème **19-2**.29 est encore valable si on suppose uniquement g continue sur $\mathcal{U}$ et **localement lipschitzienne par rapport à la variable d'espace**, c'est à dire

$$\forall \mathcal{K} \text{ compact } \subset \mathcal{U} \quad \exists K > 0 \quad \forall (t, x) \text{ et } (t, y) \in \mathcal{K}$$
$$\|f(t, x) - f(t, y)\| \leqslant K \|x - y\|$$

EXERCICE 19-2.31 Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1 telle que

$$\forall t \geqslant 0 \quad f(t,0) \geqslant 0$$

Montrer que toute solution de $x'(t) = f(t, x(t))$ telle que $x(0) = x_0 \geqslant 0$ vérifie

$$\forall t \geqslant 0 \quad x(t) \geqslant 0$$

19-2.3 Exemples d'études

19-2.3.1 Utilisation d'un partitionnement du plan

Nous étudions ici le cas d'équations pour lesquelles l'espace des phases est de dimension 2 : soit il s'agit d'un système autonome de la forme

$$(E_1) : \begin{cases} x' = f(x,y) \\ y' = g(x,y) \end{cases}$$

où la fonction inconnue[19] $t \mapsto X(t) = (x(t), y(t))$ prend ses valeurs dans $\mathbb{R}^2$, soit il s'agit d'une équation non autonome (scalaire)

$$x' = f(t,x)$$

se ramenant en fait à

$$(E_2) : \begin{cases} t' = 1 \\ x' = f(t,x) \end{cases}$$

- Dans le cas de (E_1), on essaie de tracer les courbes intégrales dans le plan rapporté à un repère $\left(O, \vec{I}, \vec{J}\right)$: par tout point $M_0(x_0, y_0)$ (pour lequel $(x_0, y_0) \in \mathcal{U}$) passe une unique trajectoire maximale de (E_1), qui est tangente en M_0 au vecteur

$$\vec{V}(M_0) = f(x_0, y_0)\,\vec{I} + g(x_0, y_0)\,\vec{J}$$

(si ce vecteur est non nul). Les signes de $f(M_0)$ et $g(M_0)$ permettent d'obtenir des informations sur la monotonie des fonctions

$$t \mapsto x(t) \text{ et } t \mapsto y(t)$$

au voisinage de t_0 correspondant à M_0, le rapport $\dfrac{g(M_0)}{f(M_0)}$ représentant (si $f(M_0) \neq 0$) la pente de la tangente à la trajectoire en M_0. On détermine donc les "courbes"

$$\Gamma_x = \{M(x,y) \mid (x,y) \in \mathcal{U} \text{ et } f(x,y) = 0\}$$

[19] Il peut s'agir de $X(t) = (x(t), x'(t))$ pour un système du premier ordre équivalent à une équation autonome du second ordre

$$x'' = h(x, x')$$

$$\Gamma_y = \{M(x,y) \mid (x,y) \in \mathcal{U} \text{ et } g(x,y) = 0\}$$

Γ_x est l'ensemble des points du plan où le champ de vecteurs est colinéaire à $\vec{J}$, et où les trajectoires ont une tangente verticale. De même, Γ_y contient les points où cette tangente est horizontale. L'intersection $\Gamma_x \cap \Gamma_y$ est constitué des points d'équilibre, c'est à dire des trajectoires qui sont réduites à un point. Les courbes Γ_x et Γ_y délimitent en général des portions du plan où f et g gardent des signes constants, ce qui permet de voir le sens de l'évolution du point $M(t) = M(x(t), y(t))$ tant qu'il reste dans une telle portion. Une étude qualitative précise permet alors parfois d'obtenir plus d'information sur le comportement de l'application $t \mapsto M(t)$. Par exemple, si on peut trouver un domaine $\mathcal{D}$ du plan dans lequel le champ est "rentrant", on pourra prouver qu'une trajectoire qui rentre dans $\mathcal{D}$ y reste piégée. Si de plus $\mathcal{D}$ est compact, on pourra appliquer le résultat de l'exercice **19-2**.26 pour prouver que la solution maximale correpondante est définie sur un intervalle de la forme $]a, +\infty[$. L'exercice **19-2**.34 est un exemple classique d'étude à la fois "qualitative" et "analytique" d'une équation en dimension égale à 2.

- Dans le cas de (E_2), la situation est plus simple, puisque t' est constamment égal à 1. Si $M_0(t_0, x_0) \in \mathcal{U}$, la tangente à l'unique trajectoire maximale passant par M_0 a pour pente $m_0 = f(t_0, x_0)$. Les points où m_0 s'annule sont les points où les courbes intégrales ont une tangente horizontale :

EXEMPLE 19-2.32 Nous allons décrire les courbes intégrales correspondant aux solutions maximales de l'équation différentielle

$$(E) : x'(t) = \cos(x(t) - t)$$

Remarquons d'abord que, la fonction $(t, x) \to g(t, x) = \cos(x - t)$ est de classe C^1 sur $\mathbb{R}^2$, et donc le théorème de Cauchy-Lipschitz s'applique. De plus, si x est une solution définie sur un intervalle $I =]a, b[$ avec $b < +\infty$, on a

$$\forall t \in I \quad |x'(t)| \leqslant 1$$

et donc x possède une limite finie en b. Il est alors possible de prolonger x en une solution au delà de b. Une étude analogue est possible en a si $a > -\infty$. Les solutions maximales sont donc définies sur $\mathbb{R}$.

Le champ de vecteurs associé à (E) est

$$\vec{V}(t, x) = \vec{I} + \cos(x - t)\, \vec{J}$$

et est constant le long des droites (parallèles à la première bissectrice) d'équations

$$\mathcal{D}_\alpha : x = t + \alpha \quad (\alpha \in \mathbb{R})$$

Par chacun des points d'une telle droite passe une unique trajectoire maximale, avec une tangente ayant pour pente $m = \cos\alpha$. Il est donc géométriquement évident qu'en translatant une courbe intégrale d'un vecteur $\vec{U} = \lambda\left(\vec{I} + \vec{J}\right)$, donc dans la direction de la première bissectrice, on retrouve une courbe intégrale.

Ceci traduit simplement le fait que, si $t \mapsto x(t)$ est $\mathbb{R}$-solution de (E), la fonction $t \mapsto \lambda + x(t-\lambda)$ est aussi solution de (E). Les courbes intégrales possèdent une autre propriété d'invariance : si x est une $\mathbb{R}$-solution, il en est de même de $t \mapsto x(t+2\pi)$. La famille des courbes intégrales est donc également invariante par toute translation de vecteur $\vec{U} = 2k\pi\,\vec{I}$, avec $k \in \mathbb{Z}$. Nous nous limiterons ainsi à l'étude d'une solution maximale vérifiant une condition initiale $x(t_0) = 0$ avec $t_0 \in [0, 2\pi]$, condition à laquelle on peut toujours se ramener en combinant deux translations décrites précédemment.

Si $t_0 = 0$ (ou 2π), le point $M(t_0, x_0)$ appartient à une droite $\mathcal{D}_o$ dirigée par le vecteur $\vec{I} + \vec{J} = \vec{V}(M_0)$. En tout point M de cette droite, on a $\vec{V}(M) = \vec{V}(M_0)$, et il est donc évident que cette droite est trajectoire maximale pour (E). On vérifie d'ailleurs immédiatement que

$$t \mapsto x(t) = t \quad \text{ou } t \mapsto x(t) = t - 2\pi \text{ sont deux } \mathbb{R}\text{-solutions de } (E)$$

Le théorème de Cauchy-Lipschitz montre que deux trajectoires maximales ne peuvent se couper sans être égales. Il en résulte que, pour une condition initiale

$$x(t_0) = 0 \quad \text{avec} \quad t_0 \in]0, 2\pi[$$

nous aurons (grâce au théorème des valeurs intermédiaires)

$$\forall t \in \mathbb{R} \quad t - 2\pi < x(t) < t$$

La trajectoire reste donc entièrement dans une bande du plan délimitée par deux parallèles à la première bissectrice. Comme

$$\forall t \in \mathbb{R} \quad |x'(t)| \leqslant 1$$

la fonction $t \mapsto t - x(t) \in]0, 2\pi[$ est croissante bornée, et possède donc une limite en $\pm\infty$. Si $\lambda = \lim_{t\to+\infty} x(t)$ était inférieure à 2π, on aurait

$$\lim_{t\to+\infty} 1 - x'(t) = 1 - \cos(\lambda) > 0$$

Ceci est en contradiction avec le fait que $t \mapsto t - x(t)$ soit bornée. L'étude en $-\infty$ étant analogue, nous avons donc

$$\lim_{t\to-\infty} t - x(t) = 0 \quad \text{et} \quad \lim_{t\to+\infty} t - x(t) = 2\pi$$

ce qui montre que les trajectoires rectilignes correspondant aux solutions $t \mapsto t$ et $t \mapsto t - 2\pi$ sont asymptotes à la trajectoire maximale étudiée. En représentant le champ de vecteurs en certains points, on peut ainsi avoir une idée de la forme des courbes intégrales (figure **19.11** page 932).

EXERCICE 19-2.33 En faisant le changement de fonction inconnue

$$z(t) = t - x(t)$$

donner une expression analytique des solutions, en utilisant les résultats de la section **19-2.1.3**. Retrouver les résultats de l'étude qualitative précédente.

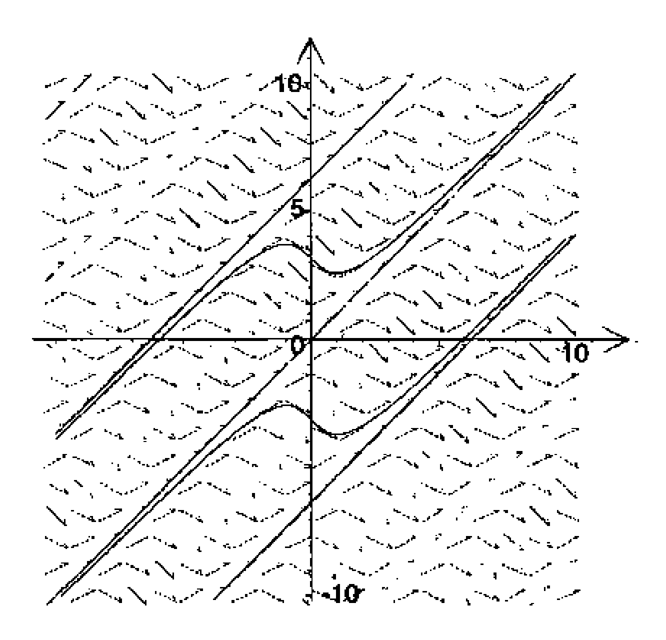


Figure **19.11** – *Champ de vecteurs et trajectoires pour* $x' = \cos(x-t)$

EXERCICE 19-2.34 On considère le système autonome d'ordre 2

$$(E) : \begin{cases} x' = x(-2+3y) \\ y' = y(3-x) \end{cases}$$

et on étudie une solution maximale correspondant à une condition initiale

$$(x(0), y(0)) = (x_0, y_0) \in \mathbb{R}^{*+} \times \mathbb{R}^{*+}$$

- Partitionner le quart de plan $\{x \geqslant 0, y \geqslant 0\}$ selon le signe de x' et y', et montrer que la solution maximale considérée vérifie

$$\forall t \in \mathbb{R} \quad x(t) > 0 \text{ et } y(t) > 0$$

- Déterminer les points d'équilibre.

- On suppose que $x_0 < 3$ et $y_0 \geqslant \dfrac{2}{3}$. Montrer qu'on peut trouver un $\alpha > 0$ tel que les fonctions x et y soient définies et strictement croissantes sur $[0, \alpha]$, et montrer ensuite qu'on peut choisir α pour avoir $x(\alpha) = 3$. Poursuivre l'étude au delà de la valeur α.

- Comme les solutions de (E) ne s'annulent pas, le système (E) entraine, si I est l'intervalle de définition de la solution maximale

$$\forall t \in I \quad \frac{3 - x(t)}{x(t)} x'(t) = \frac{3y(t) - 2}{y(t)} y'(t)$$

(on a "séparé les variables"). Montrer qu'une trajectoire maximale de (E) est incluse dans une ligne de niveau

$$C_\lambda = \{M(x,y) \mid x > 0,\ y > 0 \text{ et } f(x,y) = \lambda\}$$

pour une fonction $f : \mathbb{R}^{*+} \times \mathbb{R}^{*+} \to \mathbb{R}$ que l'on déterminera. En déduire que ces solutions maximales sont définies sur $\mathbb{R}$ et sont toutes périodiques (voir figure **19.12** page 933).

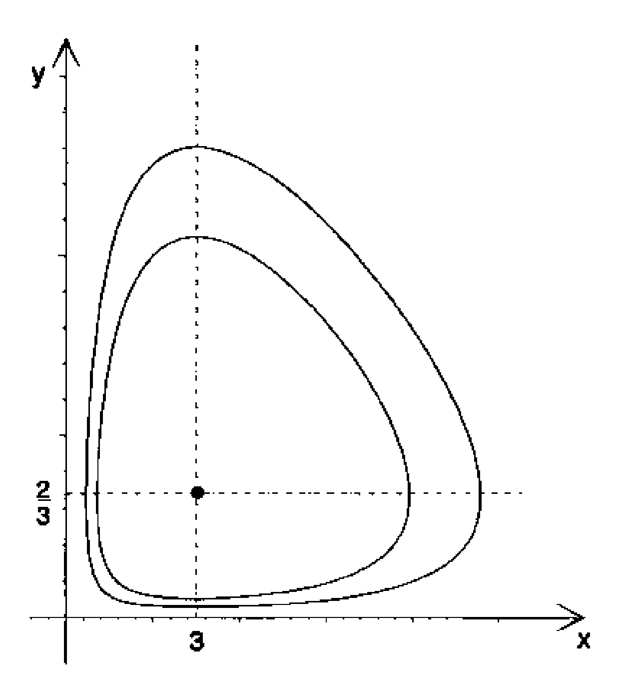


Figure **19.12** – *Deux trajectoires du système différentiel autonome* $x' = x(3y-2) \quad y' = y(3-x)$.

19-2.3.2 Equations à variables séparables

Il s'agit d'équations scalaires pouvant s'écrire sous la forme

$$(E) : a(x)\,x' = b(t)$$

où a et b sont des fonctions continues définies respectivement sur des intervalles K et I de $\mathbb{R}$ (I ouvert). Les solutions de (E) sont donc des fonctions dérivables définies sur un sous-intervalle $J \subset I$ à valeurs dans K. **Si on cherche d'abord des solutions à valeurs dans un intervalle ouvert $K_0 \subset K$ sur lequel a ne s'annule pas,** l'équation peut s'écrire

$$x' = \frac{b(t)}{a(x)}$$

(équation "résolue en x'") et le théorème de Cauchy-Lipschitz s'applique si a est de classe C^1 sur K_0 : il suffit de remarquer que la fonction

$$g : I \times K_0 \to \mathbb{R} \quad (t,x) \mapsto \frac{b(t)}{a(x)}$$

est alors localement lipschitzienne par rapport à la variable x. Ensuite on étudie les éventuels raccordements en un point t_1 pour lequel une solution x à valeurs dans K_0 définie à gauche de t_1 vérifie $\lim\limits_{t \to t_1^-} x(t) = x_1$ avec $a(x_1) = 0$.

On peut souvent trouver une expression analytique des solutions, par une méthode analogue à celle de la section **19-2.1.3** (où a été traité les cas particulier des équations autonomes, qui sont à variables séparables) :

Si on cherche une solution de (E) définie sur un intervalle J contenant t_0 et vérifiant $x(t_0) = x_0 \in K$, nous aurons

$$\forall t \in J \quad \int_{t_0}^{t} a(x(u))\, x'(u)\, du = \int_{t_0}^{t} b(u)\, du$$

Si A et B représentent respectivement des primitives des fonctions continues a et b sur K et I, on aura

$$\forall t \in J \quad A(x(t)) - A(x_0) = B(t) - B(t_0)$$

Lorsque x prend ses valeurs dans un intervalle où a garde un signe constant, A est strictement monotone sur cet intervalle et peut alors être inversée, ce qui permettra d'expliciter $x(t)$ à l'aide de l'application réciproque A^{-1}.

REMARQUE 19-2.35 Pour une équation de la forme

$$x'(t) = \alpha(x(t))\, \beta(t)$$

avec $\alpha \in C^1(K, \mathbb{R})$ et $\beta \in C^0(I, \mathbb{R})$, on invoquera le théorème de Cauchy-Lipschitz avant de séparer les variables :

Résolvons par exemple, dans un intervalle inclus dans $\mathbb{R}^{*+}$ l'équation

$$(E) : tx'(t) = 1 - x^2(t)$$

Plus précisément, cherchons à déterminer les solutions maximales vérifiant la condition initiale $x(t_0) = x_0 \in \mathbb{R}$ avec $t_0 > 0$. Comme

$$g : (t, x) \mapsto \frac{1 - x^2}{t}$$

est de classe C^1 sur l'ouvert $\mathcal{U} = \mathbb{R}^{*+} \times \mathbb{R}$, le théorème de Cauchy-Lipschitz nous assure l'existence et l'unicité de cette solution maximale, définie sur un intervalle ouvert I tel que $t_0 \in I \subset \mathbb{R}^{*+}$ (l'exercice qui suit nous montrera que ce n'est plus vrai si on cherche à résoudre sur un intervalle contenant 0, puisque (E) n'est pas résolue en x' : toute solution définie en 0 doit nécessairement vérifier $x(0) = \pm 1$).

Si $x_0 = \pm 1$, on dispose d'une solution évidente définie sur $]0, +\infty[$, constante égale à x_0. C'est donc la solution maximale cherchée. Si $x_0 \neq \pm 1$, on a forcément

$$\forall t \in I \quad x^2(t) \neq 1$$

puisque deux trajectoires distinctes ne peuvent se couper. On pourra alors séparer les variables et écrire

$$\forall t \in I \quad \frac{x'(t)}{1 - x^2(t)} = \frac{1}{t}$$

EXERCICE 19-2.36 On poursuit l'étude qui précède : si $x_0 \in]-1, +1[$, montrer que x est strictement croissante sur I, alors qu'il y a décroissance stricte si $|x_0| > 1$. Montrer qu'il existe $\varepsilon \in \{\pm 1\}$ et $t_1 > 0$ (que l'on explicitera en fonction de t_0 et x_0) tel que

$$\forall t \in I \quad \frac{1 + x(t)}{1 - x(t)} = \varepsilon \frac{t^2}{t_1^2}$$

ce qui donne

$$\forall t \in I \quad x(t) = 1 - \frac{2t_1^2}{t_1^2 + \varepsilon t^2}$$

En déduire que la solution maximale est définie sur $]0, +\infty[$ si $|x_0| < 1$, sur $]t_1, +\infty[$ si $x_0 > 1$ et $]0, t_1[$ si $x_0 < -1$.

On cherche à présent les courbes intégrales de (E) non nécessairement inscrites dans le demi-plan $t > 0$. Montrer d'abord qu'une transformation géométrique simple permet d'obtenir les trajectoires incluses dans le demi-plan $t < 0$. Déterminer les trajectoires maximales passant par les points de coordonnées $(0, 1)$ et $(0, -1)$.

19-2.3.3 Equations homogènes

Une équation homogène (résolue en x') est une équation de la forme

$$(E) : x'(t) = f\left(\frac{x(t)}{t}\right)$$

où f est une fonction de classe $\mathcal{C}^1$ sur un intervalle ouvert $K \subset \mathbb{R}$. La forme même de l'équation nous amène à rechercher des trajectoires dans le cône ouvert (de sommet O)

$$\mathcal{U} = \{t(1, x_0), \quad t > 0 \text{ et } x_0 \in K\}$$

(pour les trajectoires incluses dans le demi-plan $t < 0$, on remarquera que la famille des courbes intégrales est invariante par symétrie par rapport à l'origine : si $t \mapsto x(t)$ est I-solution de (E), l'application $t \mapsto -x(-t)$ est $-I$-solution). Le champ de vecteurs associé à (E) est constant le long des demi-droites issues de l'origine :

$$\forall x_0 \in K \quad \forall t \neq 0 \quad \vec{V}(1, x_0) = \vec{I} + f(x_0)\, \vec{J} = \vec{V}(t, t\, x_0)$$

Il est donc géométriquement évident que toute homothétie de centre O laisse invariante la famille des courbes intégrales de (E) : on traduit ainsi le fait que, si $t \mapsto x(t)$ est I-solution (avec $I \subset \mathbb{R}^{*+}$) et si $\lambda \in \mathbb{R}^*$, l'application $t \mapsto \lambda\, x\left(\dfrac{t}{\lambda}\right)$ est λI-solution.

Pour trouver une expression analytique de ces solutions, on utilise un changement de fonction inconnue, en considérant l'application

$$u : I \to K \quad t \mapsto u(t) = \frac{x(t)}{t}$$

soit

$$\forall t \in I \quad x(t) = t\, u(t)$$

Comme x, la fonction u est dérivable sur I et donc

$$\forall t \in I \quad x'(t) = u(t) + t\, u'(t) = f(u(t))$$

soit

$$(E_1) : t\,u'(t) = f(u(t)) - u(t)$$

La fonction u vérifie alors une équation aux variables séparables, que l'on intègre en suivant la méthode vue à la section précédente : si on cherche à résoudre (E) pour la condition initiale $x(t_0) = x_0$ (avec $t_0 > 0$ et $u_0 = \dfrac{x_0}{t_0} \in K$), on distinguera les cas $f(u_0) = u_0$ et $f(u_0) \neq u_0$:

- Si $f(u_0) = u_0$, la solution maximale de (E_1) est

$$\forall t > 0 \quad u(t) = u_0$$

ce qui donne la solution maximale de (E_1)

$$\forall t > 0 \quad x(t) = u_0 t$$

La trajectoire correspondante est une demi-droite issue de l'origine, de pente u_0. Ce résultat est géométriquement évident, puisqu'en tout point M de cette demi-droite, le champ de vecteur associé à (E) vaut

$$\vec{V}(M) = \vec{I} + u_0\vec{J}$$

et dirige la demi-droite.

- Si $f(u_0) \neq u_0$, on sait que

$$\forall t \in I \quad f(u(t)) - u(t) \neq 0$$

On peut alors intégrer (E_1) en séparant les variables :

$$\forall t \in I \quad \frac{u'(t)}{f(u(t)) - u(t)} = \frac{1}{t}$$

Si F est une primitive de l'application $u \mapsto [f(u) - u]^{-1}$ sur l'intervalle maximal (pour l'inclusion) $K_1 \subset K$ contenant u_0 sur lequel cette fonction ne s'annule pas, nous aurons

$$(*) \quad \forall t \in I \quad F(u(t)) - F(u_0) = \ln \frac{t}{t_0}$$

Lorsqu'on peut inverser la fonction F, on en déduira $u(t)$ en fonction de t, puis également $x(t) = t\,u(t)$. On préfère souvent écrire $(*)$ sous la forme

$$\begin{cases} t = Ce^{F(u(t))} \\ x(t) = tu(t) = Cu(t)\,e^{F(u(t))} \end{cases} \quad \text{avec } C \text{ constante réelle}$$

ce qui donne une représentation paramétrique de la trajectoire considérée, à l'aide du paramètre $u \in K_1$ (dont l'interprétation géométrique est simple). Une modification de la constante C correspond à une transformation de la courbe intégrale par homothétie de centre O.

EXERCICE 19-2.37 Décrire les courbes intégrales pour les équations différentielles

$$(E_1) : x'(t) = \frac{x(t)}{t} + \frac{t}{x(t)} + 1$$

$$(E_2) : (x(t) - t)\,x'(t) + x(t) = 0$$

19-2.3.4 Equations "aux différentielles totales"

Il s'agit d'équations de la forme

$$(E) : x'(t)\, Q(t, x(t)) + P(t, x(t)) = 0$$

où P et Q sont des fonctions continues sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, telles qu'il existe une fonction $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R})$

$$f : \mathcal{U} \ni (t, x) \mapsto f(t, x)$$

vérifiant, dans l'ouvert $\mathcal{U}$,

$$(*) \quad P = \frac{\partial f}{\partial t} \quad \text{et} \quad Q = \frac{\partial f}{\partial x}$$

On a alors

$$\forall M \in \mathcal{U} \quad df_M = P(M)\, dt + Q(M)\, dx$$

Par le théorème de dérivation d'une fonction composée, nous savons que (E) peut s'écrire

$$\frac{d}{dt}(f(t, x(t))) = \frac{\partial f}{\partial t}(t, x(t)) + x'(t)\, \frac{\partial f}{\partial x}(t, x(t)) = 0$$

et donc, si $t \mapsto x(t)$ est I-solution vérifiant une condition initiale $x(t_0) = x_0$, la fonction $t \mapsto f(t, x(t))$ est constante. La trajectoire correspondante est incluse dans la ligne de niveau

$$\Gamma_0 = \{M(t, x) \mid (t, x) \in \mathcal{U} \quad \text{et} \quad f(t, x) = f(t_0, x_0)\}$$

Le théorème des fonctions implicites nous assure que Γ_0 est, au voisinage de $M_0(t_0, x_0)$, représentation graphique d'une fonction $t \mapsto x(t)$ de classe $\mathcal{C}^1$ (qui sera solution de (E)), pour peu que l'on suppose $Q(t_0, x_0) \neq 0$. C'est d'ailleurs cette condition que l'on est amenée à considérer si l'on veut résoudre (E) en x', c'est à dire pouvoir l'écrire

$$x'(t) = -\frac{P(t, x(t))}{Q(t, x(t))}$$

C'est aux équations "résolues" en x' qu'on peut espérer appliquer le théorème de Cauchy-Lipschitz.

REMARQUE 19-2.38 Si les fonctions P et Q sont de classe $\mathcal{C}^1$ dans l'ouvert $\mathcal{U}$, une condition nécessaire d'existence de $f \in \mathcal{C}^2(\mathcal{U}, R)$ avec

$$df = Pdt + Qdx$$

est que l'on doit avoir, dans $\mathcal{U}$,

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial t}$$

(à cause du théorème de Schwarz). Nous verrons (voir théorème **23-1**.19) que cette condition est suffisante pourvu que l'ouvert $\mathcal{U}$ soit étoilé par rapport à un de ses points.

EXERCICE 19-2.39 Résoudre l'équation différentielle

$$(2x + 3t)\,x' + 3x - \cos 2t = 0$$

REMARQUE 19-2.40 Lorsqu'il n'est pas possible de trouver une fonction f vérifiant les conditions $(*)$, on peut parfois multiplier l'équation (E) par une fonction g des variables (t, x), de telle manière que l'on puisse trouver f avec

$$df = gPdt + gQdx$$

Une telle fonction g est appelée "facteur intégrant" pour (E). On peut alors résoudre l'équation "multipliée par g". Il faut prendre garde cependant que l'on risque ainsi de trouver des solutions ne convenant pas pour l'équation de départ, notamment toute fonction qui serait définie implicitement par l'équation

$$g(t, x) = 0$$

EXERCICE 19-2.41 Résoudre l'équation

$$\left(t^2 + x^2 + 2t\right) + 2xx' = 0$$

par un changement de fonction inconnue (pour se ramener à une équation linéaire), ou en utilisant un facteur intégrant ne dépendant que de t.

19-2.3.5 Equations se ramenant à des équations linéaires

Certaines équations peuvent, par changement de fonction inconnue, se ramener à des équations linéaires. Nous traiterons ici le cas des équations de Bernoulli, équations scalaires de la forme

$$(E) : x'(t) = a(t)\,x(t) + b(t)\,(x(t))^{\alpha}$$

où $a, b \in C^0(I, \mathbb{R})$ et $\alpha \in \mathbb{R}$ sont donnés. Si $\alpha = 0$ ou 1, il s'agit d'une équation linéaire. Si α n'est pas un entier, on cherchera des solutions à valeurs $\mathbb{R}^{*+}$ (pour pouvoir donner un sens à x^{α}) et on pourra appliquer le théorème de Cauchy-Lipschitz, puisque l'application

$$g : I \times \mathbb{R}^{*+} \to \mathbb{R} \quad (t, x) \mapsto g(t, x) = a(t)\,x + b(t)\,x^{\alpha}$$

est alors localement lipschitzienne par rapport à la variable x. Si $\alpha \geqslant 2$ est entier, il n'y a pas de restriction *a priori* sur l'ensemble des valeurs prises par une solution de (E). Mais, comme le théorème de Cauchy-Lipschitz s'applique dans ce cas, on peut remarquer que, si une solution maximale s'annule en un point t_0, c'est forcément la solution identiquement nulle. On cherchera donc aussi, dans le cas où $\alpha \in \mathbb{N}$, des solutions ne s'annulant pas, donc à valeurs dans $\mathbb{R}^{*+}$ ou $\mathbb{R}^{*-}$.

En écrivant (E) sous la forme

$$\frac{x'(t)}{x^{\alpha}(t)} = a(t)\,\frac{1}{x^{\alpha-1}(t)} + b(t)$$

on se ramène à une équation linéaire si on effectue le changement de fonction inconnue

$$z(t) = \frac{1}{x^{\alpha-1}(t)}$$

On obtient alors

$$(E_1) : \frac{1}{1-\alpha} z'(t) = a(t) z(t) + b(t)$$

La théorie des équations linéaires nous montre que les solutions maximales de (E_1) sont définies sur I tout entier. Si z est une telle solution, on reviendra ensuite à une solution de (E) en trouvant les sous-intervalles J de I sur lesquels $z > 0$ (ou z ne s'annule pas si α est un entier pair) et en déterminant sur J les fonctions x vérifiant

$$x^{\alpha-1}(t) = \frac{1}{z(t)}$$

EXERCICE 19-2.42 Résoudre sur $]0, +\infty[$ l'équation

$$t\, x'(t) - x(t) = t\, x^2(t)$$

Déterminer la solution maximale vérifiant une condition initiale $x(t_0) = x_0$ pour $t_0 > 0$ et $x_0 \in \mathbb{R}$.

19-3 Exercices

EXERCICE 19-3.1 Trouver $f \in \mathcal{C}^0(\mathbf{R}, \mathbf{R})$ telle que

$$\forall x \in \mathbf{R} \quad f(x) = 1 + \int_{-x}^{0} (t-x) f(x+t) dt$$

EXERCICE 19-3.2 Soient a et b deux fonctions continues de $\mathbf{R}$ dans $\mathbb{C}$ 1-périodiques. On considère l'équation différentielle

$$(E) \quad y' + a(t) y = b(t)$$

Si f est une solution de l'équation homogène associée, trouver une relation entre $f(t)$ et $f(t+1)$. En déduire que (E) n'a, en général, qu'une seule solution 1-périodique.

EXERCICE 19-3.3 Quelle condition doit vérifier l'endomorphisme u de $\mathbb{C}^n$ pour que toutes les solutions de $y' = uy$ soient bornées sur $\mathbf{R}$? sur $[0, +\infty[$?

EXERCICE 19-3.4 Résoudre : $\left\{ \begin{array}{lll} x' & = & tx - y + t\cos t - t^3 \sin t \\ y' & = & x + ty + t \sin t + t^3 \cos t \end{array} \right.$

EXERCICE 19-3.5 Soit $A : \mathbf{R} \to \mathcal{M}_2(\mathbb{C})$ une application continue et Y_1, Y_2 deux solutions indépendantes de $Y' = AY$. On suppose que A est 2π-périodique. Montrer qu'il existe $B \in \mathcal{M}_2(\mathbb{C})$ telle que l'application

$$\psi : t \to [Y_1(t), Y_2(t)] \exp(tB)$$

soit 2π-périodique.

EXERCICE 19-3.6 Soit u une application continue de $\mathbb{R}_+$ dans $\mathbb{R}$ et f une application continue de $\mathbb{R}_+$ dans $\mathbb{R}_+$. On suppose qu'il existe une constante A telle que, pour tout $x \in \mathbb{R}_+$ on ait :

$$u(x) \leqslant A + \int_0^x f(t)u(t)dt$$

Montrer que

$$u(x) \leqslant A \exp\left(\int_0^x f(t)dt\right)$$

Soit (E) l'équation différentielle $y'' + y(1 + g(t)) = 0$ où g est une application continue de $\mathbb{R}_+$ dans $\mathbb{R}$ intégrable. Montrer que toute solution de (E) est bornée.

EXERCICE 19-3.7 Résoudre (les paramètres a, b, c, k sont réels) :

$$\left\{\begin{array}{rcrcr} x' &=& bz &-& cy \\ y' &=& cx &-& az \\ z' &=& ay &-& bx \end{array}\right. \quad \left\{\begin{array}{rcl} y_1' &=& k^2(-2y_1 + y_2) \\ y_2' &=& k^2(y_1 - 2y_2 + y_3) \\ y_{n-1}' &=& k^2(y_{n-2} - 2y_{n-1} + y_n) \\ y_n' &=& k^2(y_{n-1} - 2y_n) \end{array}\right.$$

$$\left\{\begin{array}{rcl} x' &=& -2x + 6y - 2z \\ y' &=& -x + 3y - z - 2t \\ z' &=& -x + 3y + z - 5\sin t \end{array}\right.$$

EXERCICE 19-3.8 Soit q la somme d'une série entière de rayon de convergence R, et $a, b \in \mathbb{R}$. Montrer que la solution u de $y'' - q(x)y = 0$ avec $u(0) = a$ et $u'(0) = b$ est somme d'une série entière de rayon de convergence $R' \geqslant R$.

EXERCICE 19-3.9 Soit $\alpha > 0$ et l'équation (E) : $x'' + \frac{1}{t}x' + (1 - \frac{\alpha^2}{t^2})x = 0$.

1. Lorsque $\alpha = \frac{1}{2}$, résoudre (E).
2. Dans le cas général, montrer que (E) possède sur $]0, +\infty[$ une unique solution J_α équivalente en 0 à t^α et que toute solution non proportionnelle à J_α est infiniment grande au voisinage de 0.

EXERCICE 19-3.10 Si $f \in \mathcal{C}^0(\mathbb{R}_+, \mathbb{R})$ avec f intégrable, les solutions sur $[0, +\infty[$ de

$$y'' + fy = 0$$

peuvent elles être toutes bornées ? Si f est solution bornée de $y'' + e^{-t^2}y = 0$ avec f^2 intégrable sur $\mathbb{R}^+$, montrer que f' est bornée puis que $f \to 0$ en $+\infty$.

EXERCICE 19-3.11 Soit I un intervalle de $\mathbb{R}$ et a, b deux applications continues de I dans $\mathbb{R}$. On considère l'équation différentielle

$$(E) \quad y'' + ay' + by = 0$$

Soit φ une solution non identiquement nulle de (E).

1. Montrer que, pour K sous intervalle compact de I l'ensemble

$$Z_\varphi = \{t \in K \mid \varphi(t) = 0\}$$

est fini. Que peut-on en déduire pour l'ensemble des zéros de φ sur I ?

2. On suppose que φ s'annule au moins deux fois sur I, et on considère α, β deux zéros consécutifs de φ (expliquer pourquoi cela a un sens). Montrer que, si ψ est une I-solution de (E) non proportionnelle à φ, alors $\psi(\alpha)\psi(\beta) \neq 0$ et ψ s'annule une seule fois sur $]a, b[$.

3. On considère deux équations différentielles :

$$(E_f) \quad y'' + fy = 0 \text{ et } (E_g) \quad y'' + gy = 0$$

avec f et g continues de I dans $\mathbb{R}$ et $f \geqslant g$. Montrer que, si u est solution de E_g ayant α et β pour zéros consécutifs, et v est solution de E_f ne s'annulant pas simultanément en α et β, alors v admet un zéro dans $]\alpha, \beta[$.

EXERCICE 19-3.12 Soit $f \in \mathcal{C}^2(\mathbf{R}, \mathbf{R})$ avec $f''(t) + f'(t) + f(t) \sim t$ pour $t \to +\infty$. Donner un équivalent de $f(t)$ au voisinage de $+\infty$.

EXERCICE 19-3.13 Soit $P \in \mathbb{R}[X]$ tel que $\forall x \in \mathbb{R} \quad P(x) \geqslant 0$. Montrer que le polynôme $Q(X) = \sum_{n=0}^{+\infty} P^{(n)}(X)$ possède la même propriété.

EXERCICE 19-3.14 Soit f développable en série entière sur $]-R, R[$: $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. On considère l'équation différentielle

$$(\mathcal{E}) \quad \begin{cases} y' - y = f\left(\dfrac{1}{x}\right) \\ x > \dfrac{1}{R} \end{cases}$$

- Montrer que $(\mathcal{E})$ possède une seule solution bornée au voisinage de $+\infty$. On la note y.
- Soit $\psi_n(x) = -\int_x^{+\infty} \frac{e^{-t}}{t^n} dt$. Montrer que $y(x) = e^{-x} \sum_{n=0}^{+\infty} a_n \psi_n(x)$.
- Donner un équivalent de ψ_n en $+\infty$.
- Montrer que y est équivalent à $-f\left(\dfrac{1}{x}\right)$ en $+\infty$.

EXERCICE 19-3.15 Soit f une fraction rationnelle réelle et a son plus grand pôle réel. Dans $]a, +\infty[$ on considère l'équation différentielle $(\mathcal{E})$ $y' - y = f(x)$.

- Montrer qu'il existe une unisue solution de $(\mathcal{E})$ notée y_f telle que $y_f(x) = o(e^x)$ pour $x \to +\infty$.
- Donner une condition nécessaire et suffisante sur f pour que $\lim_{+\infty} y_f = 0$.
- Si la condition précédente n'est pas vérifiée, montrer qu'il existe un polynôme P et une fonction ε de limite nulle en ∞ telle que $y_f(x) = P(x) + \varepsilon(x)$. P est-il unique ?
- Montrer que, si g est une fraction rationnelle réelle, avec $g(x) = o\Big(f(x)\Big)$ en $+\infty$, alors $y_g(x) = o\Big(y_f(x)\Big)$ pour $x \to +\infty$.

- Montrer que $y_f(x) \sim -f(x)$ en $+\infty$.

EXERCICE 19-3.16 Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $M \in C^1(\mathbb{R}, \mathcal{M}_n(\mathbb{C}))$ solution de l'équation différentielle

$$\frac{dM}{dt} = AM - MA$$

Que peut-on dire du spectre de M ?

EXERCICE 19-3.17 A quelle relation doivent satisfaire les fonctions numériques p et q pour que l'équation différentielle $y'' + p(x)y' + q(x)y = 0$ admette deux solutions particulières u et v vérifiant $uv = 1$? Montrer que l'on obtiendrait la même condition en imposant la condition $u^2 - v^2 = 1$ ou $u^2 + v^2 = 1$. Ces solutions u et v peuvent-elles être réelles ? Application : résoudre $(1 + \cos 4x)y'' - 2y' \sin 4x - 8y = 0$.

EXERCICE 19-3.18 Soit $f : [0, a[\to \mathbb{R}$, indéfiniment dérivable et telle que

$$\forall x \in [0, a[\ \ \forall n \in \mathbb{N} \quad f^{(n)}(x) \geqslant 0$$

Montrer que $\forall x \in [0, a[\ \ f(x) = \sum_{n \geqslant 0} \frac{f^{(n)}(0)}{n!} x^n$.

Application : Montrer que sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ $x \mapsto \tan x$ est somme d'une série entière. Quel est le rayon de convergence de cette série ?

Si g est la solution maximale de l'équation différentielle $y' = x + y^2$ pour la condition initiale $y(0) = 0$, montrer qu'il existe un réel $a > 0$ tel que g soit définie sur $[0, a[$ avec $\lim_{x \to a^-} g(x) = +\infty$ et g est sur cet intervalle somme d'une série entière. Comment peut-on calculer les coefficients de cette série ?

EXERCICE 19-3.19 Intégrer l'équation $y' = (\frac{1}{2} + i)y - y|y|^2$

EXERCICE 19-3.20 Etudier les courbes intégrales de $y = y'^3 - 3y'$

EXERCICE 19-3.21 Montrer que les courbes intégrales de $yy' + \frac{2x^3}{y^2 + 3x^2} = 0$ sont fermées. Calculer leur aire intérieure.

EXERCICE 19-3.22 Etudier les solutions maximales de l'équation différentielle

$$y' = e^{y^2 - y} - 1$$

et notamment leurs limites aux bornes des intervalles de définition. Y a-t-il des solutions périodiques ?

EXERCICE 19-3.23 Montrer que les solutions maximales de $y'' + \sin xy = 0$ sont définies sur $\mathbb{R}$ tout entier.

EXERCICE 19-3.24 Soit f de classe ∞ de $\mathbb{R}^n$ dans lui-même. On suppose que $\forall x \in \mathbb{R}^n \ \ \|df_x\| \leqslant C$, où C est un réel positif donné. On considère une solution T-périodique non constante de l'équation différentielle $x'(t) = f(x(t))$. Minorer T.

EXERCICE 19-3.25 On considère l'équation différentielle $2y'' = 1 - 3y^2$ avec conditions initiales $y(0) = 0$ et $y'(0) = 0$. Soit f la solution maximale de ce problème. Montrer qu'il existe $a > 0$ tel que f soit croissante sur $[0, a]$ et $f'(a) = 0$; on exprimera a sous forme intégrale. Montrer ensuite que f est de période $2a$ et paire.

Chapitre 20

Arcs paramétrés

20-1 Etude affine

Dans cette section, $\mathcal{E}_n$ désigne un espace affine réel de dimension $n = 2$ ou 3, et E_n l'espace vectoriel sous-jacent. Le choix d'un point arbitraire $O \in \mathcal{E}_n$ permet d'identifier E_n et $\mathcal{E}_n$ à l'aide de l'application

$$E_n \to \mathcal{E}_n \quad \overrightarrow{u} \mapsto M = O + \overrightarrow{u}$$

L'égalité $M = O + \overrightarrow{u}$ s'écrit aussi $\overrightarrow{OM} = \overrightarrow{u}$. Le choix d'une base $\mathcal{B}$ de E_n nous donne un repère cartésien de $\mathcal{E}_n$

$$\mathcal{R} = (O, \mathcal{B})$$

Les coordonnées d'un point M dans $\mathcal{R}$ sont celles de $\overrightarrow{OM}$ dans $\mathcal{B}$: lorsque $n = 3$ et pour $\mathcal{B} = \left\{\overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right\}$, le point M a pour coordonnées $(x, y, z) \in \mathbb{R}^3$ dans ce repère si et seulement si

$$\overrightarrow{OM} = x\overrightarrow{I} + y\overrightarrow{J} + z\overrightarrow{K} \quad \text{ce que nous noterons } M\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Pour le moment, nous considérons que E_n est muni d'une norme $\| \|$ arbitraire[1]. L'espace $\mathcal{E}_n$ est alors muni de la distance associée, définie par

$$\forall A, B \in \mathcal{E}_n \quad d(A, B) = \left\|\overrightarrow{AB}\right\|$$

[1] Le cas spécifique où cette norme est euclidienne (c'est à dire associée à un produit scalaire) sera développé ultérieurement (dans ce cas, l'espace affine $\mathcal{E}_n$ sera dit également euclidien).

On peut alors définir la notion d'application de classe C^k, $k \in \mathbb{N}$, sur un intervalle I de $\mathbb{R}$ (non réduit à un point) à valeurs dans $\mathcal{E}_n$:

DÉFINITION 20-1.1 *Une application $t \mapsto M(t) \in \mathcal{E}_n$ définie sur un intervalle I de $\mathbb{R}$ est de classe C^k ($k \geqslant 0$) si et seulement si l'application*

$$\overrightarrow{F} : I \to E_n \quad t \mapsto \overrightarrow{F}(t) = \overrightarrow{OM(t)}$$

est de classe C^k, pour le choix d'un point O de $\mathcal{E}_n$. Cette propriété ne dépend pas du point O.

En effet, si O_1 est un autre point de $\mathcal{E}_n$, les fonctions définies sur I par

$$\overrightarrow{F}(t) = \overrightarrow{OM(t)} \text{ et } \overrightarrow{G}(t) = \overrightarrow{O_1M(t)}$$

ne diffèrent que du vecteur constant $\overrightarrow{OO_1}$. En particulier, si $k \in \mathbb{N}^*$ et si les applications sont de classe C^k, nous aurons

$$\forall t \in I \quad \forall i \in \{1, \dots, k\} \quad \overrightarrow{F}^{(i)}(t) = \frac{d^i \overrightarrow{F}}{dt^i}(t) = \overrightarrow{G}^{(i)}(t) = \frac{d^i \overrightarrow{G}}{dt^i}(t)$$

Nous noterons simplement, s'il n'y a pas d'ambigüité

$$\overrightarrow{F}^{(i)}(t) = \overrightarrow{M}^{(i)}(t)$$

Si $\mathcal{R} = (O, \mathcal{B})$ est un repère cartésien d'origine O, les coordonnées de $\overrightarrow{F}(t)$ dans $\mathcal{B}$ sont les coordonnées de $M(t)$ dans $\mathcal{R}$. Il en résulte que l'application $t \mapsto M(t)$ est de classe C^k sur I si et seulement si la fonction $I \to \mathbb{R}^n$ définie par

$$t \mapsto \text{ le } n\text{-uplet des coordonnées de } M(t) \text{ dans } \mathcal{R}$$

est de classe C^k. Avec $n = 3$

$$t \mapsto M(t) \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \text{ est } C^k \Leftrightarrow x, y \text{ et } z : I \to \mathbb{R} \text{ sont } C^k$$

et on a alors

$$\forall t \in I \quad \forall i \in \{1, \dots, k\} \quad \overrightarrow{M}^{(i)}(t) = x^{(i)}(t)\ \overrightarrow{I} + y^{(i)}(t)\ \overrightarrow{J} + z^{(i)}(t)\ \overrightarrow{K}$$

20-1.1 Définitions

DÉFINITION 20-1.2 *On appelle arc paramétré de classe C^k dans l'espace affine $\mathcal{E}_n$ toute application de classe C^k définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathcal{E}_n$.*

Si $M : I \ni t \mapsto M(t)$ est un tel arc, il ne faut pas confondre cet arc avec son image, qu'on appelle aussi son **support**

$$M(I) = \{M(t),\ t \in I\}$$

Se donner un arc paramétré, c'est se donner cet ensemble mais aussi un mode de description, de parcours de ce support : si $\mathcal{E}_2$ est un espace affine euclidien rapporté à un repère orthonormé, les arcs

$$[0,2\pi] \ni t \mapsto M(t) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad \text{et} \quad [0,1] \ni t \mapsto P(t) \begin{pmatrix} \cos 4\pi t \\ \sin 4\pi t \end{pmatrix}$$

ont même support (le cercle unité), parcouru dans les deux cas "à vitesse constante". Mais ces deux arcs sont de natures distinctes, puisque le support est parcouru deux fois dans le second cas. Par contre, il n'y a pas de grosse différence entre le premier de ces arcs et l'application

$$[0,\pi] \ni t \mapsto N(t) \begin{pmatrix} \cos 2t \\ \sin 2t \end{pmatrix}$$

Nous dirons que l'on passe de l'un à l'autre par un "changement de paramètre admissible", conformément aux définitions suivantes :

DÉFINITION 20-1.3 *Si $M : I \ni t \mapsto M(t)$ et $P : J \ni s \mapsto P(s)$ sont deux arcs paramétrés de classe C^k, on dit que ces deux arcs sont C^k-équivalents si et seulement s'il existe un difféomorphisme de classe C^k*

$$\theta : I \to J \quad t \mapsto s = \theta(t)$$

tel que $M = P \circ \theta$, soit

$$\forall t \in I \quad M(t) = P(\theta(t))$$

Il est clair que les deux arcs jouent des rôles symétriques dans cette définition, puisque

$$M = P \circ \theta \Leftrightarrow P = M \circ \theta^{-1}$$

avec $\theta^{-1} : J \to I$ qui est aussi un C^k-difféomorphisme d'intervalles. Si M et P sont deux arcs C^k avec $M = P \circ \theta$ et $\theta : I \to J$ un C^k-difféomorphisme, nous dirons que $t \mapsto s = \theta(t)$ est un changement de paramétrage admissible pour l'arc paramétré M.

Comme le composé de deux C^k-difféomorphismes en est encore un, la relation binaire que nous venons de définir est une relation d'équivalence sur l'ensemble des arcs paramétrés de classe C^k. Une classe d'équivalence pour cette relation est appelée "**arc géométrique**". Une propriété d'un arc paramétré de classe C^k sera dite géométrique si elle est invariante par changement de paramètre admissible, c'est à dire si elle est vérifiée par tout arc C^k-équivalent. Par exemple, deux arcs C^k-équivalents ont nécessairement même support, qu'on appellera support de l'arc géométrique correspondant. Dans ce qui suit, nous confondrons souvent un arc géométrique γ de classe C^k et un représentant arbitraire de γ.

DÉFINITION 20-1.4 *Si $M : I \ni t \mapsto M(t)$ est un arc C^k, et si M_0 est un point du support de cet arc, on appelle multiplicité de M_0 le cardinal de l'ensemble des **antécédents** de M_0*

$$m(M_0) = \text{card}\{t \in I \mid M(t) = M_0\}$$

($m(M_0) = +\infty$ si cet ensemble est infini). M_0 est dit simple si $m(M_0) = 1$.

Un arc est simple si tous les points de son support le sont, avec une nuance dans le cas d'un arc fermé :

DÉFINITION 20-1.5 *Si $[a,b]$ est un segment de $\mathbb{R}$ (avec $a < b$) et si*

$$M : [a,b] \ni t \mapsto M(t)$$

est un arc de classe C^k, l'origine de cet arc est le point $A = M(a)$, son extrémité est $B = M(b)$. L'arc est dit fermé si $A = B$. Dans ce cas, on dit qu'il est simple si l'application M est injective en restriction à $[a,b[$.

Si γ est un arc et M_0 est un point du support de γ, la multiplicité de M_0 est clairement une notion géométrique, indépendante du paramétrage admissible utilisé.

Lorsque γ est un arc géométrique, le choix d'un paramétrage particulier

$$M : I \ni t \mapsto M(t)$$

permet de classer les différents paramétrages de γ en deux catégories :

- Les paramétrages $P : J \ni s \mapsto P(s)$ avec $M = P \circ \theta$ pour lesquels $\theta : I \to J$ est un C^k-difféomorphisme croissant (il en est alors de même de θ^{-1}).
- Les paramétrages $N : K \ni u \mapsto N(u)$ avec $M = N \circ \psi$, avec cette fois $\psi : I \to K$ décroissant (strictement).

On dit que ces deux classes correspondent à deux orientations opposées sur l'arc γ. Un changement de paramètre admissible entre deux paramétrages d'une même classe "conserve l'orientation" de l'arc. Cette notion correspond en général à la notion de "sens de parcours" du support, notamment dans le cas d'arcs simples[2].

20-1.2 Indices fondamentaux

20-1.2.1 Définition

Soit γ un arc géométrique de classe C^k et $I \ni t \mapsto M(t)$ un représentant de γ. Pour un entier $i \in \{1, \ldots, k\}$, considérons le sous espace de E_n

$$T_i(t_0) = \text{vect}\left(\overrightarrow{M}^{(j)}(t_0)\right)_{1 \leqslant j \leqslant i}$$

Ce sous-espace est appelé $i^{\text{ème}}$ sous-espace fondamental en t_0 à l'arc paramétré M. Il s'agit d'une notion invariante par changement de paramètre admissible :

[2]Mais pour l'arc (de $\mathbb{R}^2$)

$$[-1,1] \ni t \mapsto M(t) = \left(t^2, \sin(t^2)\right)$$

le changement de paramètre admissible $t \mapsto -t$ strictement décroissant ne change pas le mode de parcours du support.

PROPOSITION 20-1.6 Si $I \ni t \mapsto M(t)$ et $P : J \ni s \mapsto P(s)$, avec $M = P \circ \theta$ sont deux arcs paramétrés équivalents, les $i^{\text{ème}}$ sous-espace fondamentaux à ces arcs respectivement en $t_0 \in I$ et $s_0 = \theta(t_0)$ sont égaux.

Démonstration : On montre par récurrence sur $i \in \{1, \dots, k\}$ qu'il existe des fonctions $(\varphi_{i,j})_{1 \leqslant j \leqslant i}$ définies sur I, avec $\varphi_{i,j}$ de classe C^{k-i} telles que

$$\forall i \in \{1, \dots, k\} \quad \overrightarrow{M}^{(i)}(t) = \left[\theta'(t)\right]^i \overrightarrow{P}^{(i)}(\theta(t)) + \sum_{j=1}^{i-1} \varphi_{i,j}(t)\, \overrightarrow{P}^{(i-j)}(\theta(t))$$

Comme θ est un C^k-difféomorphisme, on a $\theta'(t_0) \neq 0$, ce qui montre que la famille $\left(\overrightarrow{M}^{(i)}(t_0)\right)_{1 \leqslant i \leqslant k}$ est triangulaire par rapport à la famille $\left(\overrightarrow{P}^{(i)}(s_0)\right)_{1 \leqslant i \leqslant k}$. Les sous-espaces fondamentaux respectifs en t_0 et s_0 sont donc égaux. ∎

Pour effectuer une étude locale de l'arc au voisinage du paramètre t_0 on est ainsi amené à définir, lorsqu'ils existent, 2 ou 3 indices fondamentaux (selon que l'arc étudié est plan ou non) :

- Le premier indice fondamental, traditionnellement noté p, est défini par

$$p = \min\left\{i \in \{1, \dots, k\} \mid \overrightarrow{M}^{(i)}(t_0) \neq \overrightarrow{0}\right\}$$

  (lorsqu'un tel indice n'existe pas, on dit que t_0 est un point totalement stationnaire pour l'arc paramétré). Le cas le plus fréquent est celui où $p = 1$. Le point t_0 est alors dit **régulier** pour l'arc paramétré. Lorsque $p > 1$ (ou n'existe pas), t_0 est un point **stationnaire** (ou singulier).

- Le second indice, noté q, est défini par

$$q = \min\left\{i \in \{1, \dots, k\} \mid i > p \quad \text{et} \quad \overrightarrow{M}^{(i)}(t_0) \text{ indépendant de } \overrightarrow{M}^{(p)}(t_0)\right\}$$

  Le cas le plus fréquent est $p = 1$ et $q = 2$, et t_0 est alors dit point **birégulier** de l'arc. Pour un arc plan, lorsque p et q existent, la base $\left(\overrightarrow{M}^{(p)}(t_0), \overrightarrow{M}^{(q)}(t_0)\right)$ de E_2 sera commode pour faire une étude locale de l'arc au voisinage de t_0 (dans un repère d'origine $M(t_0)$).

- Enfin, pour un arc qui n'est pas plan ($n = 3$), on pourra envisager un troisième indice fondamental

$$r = \min\left\{i \in \{1, \dots, k\} \mid i > q \text{ et } \overrightarrow{M}^{(i)}(t_0), \overrightarrow{M}^{(p)}(t_0), \overrightarrow{M}^{(q)}(t_0) \text{ libres}\right\}$$

  le repère local d'étude de l'arc étant alors

$$\mathcal{R}_{t_0} = \left(M(t_0), \left\{\overrightarrow{M}^{(p)}(t_0), \overrightarrow{M}^{(q)}(t_0), \overrightarrow{M}^{(r)}(t_0)\right\}\right)$$

Lorsqu'ils sont définis, les indices fondamentaux sont invariants par changement de paramétrage admissible.

DÉFINITION 20-1.7 *Un arc de classe C^k $I \ni t \mapsto M(t)$ est dit régulier si $k \geqslant 1$ et si*

$$\forall t \in I \quad \overrightarrow{M}'(t) \neq \overrightarrow{0}$$

Il est birégulier si $k \geqslant 2$ et si tous ses points sont biréguliers, c'est à dire

$$\forall t \in I \quad \left\{\overrightarrow{M}'(t), \overrightarrow{M}''(t)\right\} \text{ est libre}$$

Dans le cas où $n = 2$ et où les indices fondamentaux p et q existent, si on note $M_0 = M(t_0)$ et $M_h = M(t_0 + h)$ pour $t_0 + h \in I$, la formule de Taylor Young en t_0, écrite à l'ordre q, donnera

$$\overrightarrow{M_0 M_h} = \overrightarrow{OM}(t_0 + h) - \overrightarrow{OM}(t_0) = \sum_{i=p}^{q} \frac{h^i}{i!} \overrightarrow{M}^{(i)}(t_0) + h^q \overrightarrow{\varepsilon}(h)$$

avec $\lim\limits_{h \to 0} \overrightarrow{\varepsilon}(h) = \overrightarrow{0}$ En décomposant ensuite chacun des vecteurs dans la base $\left(\overrightarrow{M}^{(p)}(t_0), \overrightarrow{M}^{(q)}(t_0)\right)$ et en tenant compte du fait que les coordonnées de $\overrightarrow{\varepsilon}(h)$ tendent vers 0 dans n'importe quelle base (fixée) de E_2, nous obtiendrons grâce à la définition de q :

$$\overrightarrow{M_0 M_h} \underset{h \to 0}{=} \left(\frac{h^p}{p!} + o(h^p)\right) \overrightarrow{M}^{(p)}(t_0) + \left(\frac{h^q}{q!} + o(h^q)\right) \overrightarrow{M}^{(q)}(t_0)$$

De même, dans le cas où $n = 3$ et où les trois premiers indices fondamentaux sont définis, nous aurons

$$\begin{aligned}\overrightarrow{M_0 M_h} \underset{h \to 0}{=} & \left(\frac{h^p}{p!} + o(h^p)\right) \overrightarrow{M}^{(p)}(t_0) \\ & + \left(\frac{h^q}{q!} + o(h^q)\right) \overrightarrow{M}^{(q)}(t_0) + \left(\frac{h^r}{r!} + o(h^r)\right) \overrightarrow{M}^{(r)}(t_0)\end{aligned}$$

20-1.2.2 Tangente

Nous considérons un arc de classe C^k $I \ni t \mapsto M(t)$ et un point $t_0 \in I$ où nous supposons l'existence du premier indice fondamental p. Nous avons alors, avec les notations précédentes (formule de Taylor-Young à l'ordre p),

$$\overrightarrow{M_0 M_h} = \frac{h^p}{p!} \overrightarrow{M}^{(p)}(t_0) + h^p \overrightarrow{\varepsilon_1}(h) \quad \text{avec} \quad \lim_{h \to 0} \overrightarrow{\varepsilon_1}(h) = \overrightarrow{0}$$

On a donc, pour h tendant vers 0

$$\left\|\overrightarrow{M_0 M_h}\right\| \sim \frac{|h|^p}{p!} \left\|\overrightarrow{M}^{(p)}(t_0)\right\|$$

En particulier, on peut trouver un $\alpha > 0$ tel que

$$0 < |h| < \alpha \Rightarrow M_0 \neq M_h$$

ce qui permet de définir la droite ("sécante") M_0M_h pour h "petit" non nul. Le vecteur directeur de cette droite

$$\frac{p!}{h^p}\overrightarrow{M_0M_h}$$

possède une limite égale à $\overrightarrow{M}^{(p)}(t_0)$. On dit alors que la sécante "variable" M_0M_h (passant par le point fixe M_0) possède une position limite quand h tend vers 0, qui est la droite

$$T_{t_0} = M_0 + \text{vect}\left(\overrightarrow{M}^{(p)}(t_0)\right)$$

Cette droite est appelée tangente à l'arc paramétré au point t_0. Cette notion, comme les sous-espaces fondamentaux, est invariante par changement de paramètre admissible.

EXERCICE 20-1.8 Montrer que si $\overrightarrow{u}(h)$ est un autre vecteur directeur de la sécante M_0M_h tel que la limite

$$\lim_{h\to 0}\overrightarrow{u}(h) = \overrightarrow{U}$$

existe et soit non nulle, alors $\overrightarrow{M}^{(p)}(t_0)$ et $\overrightarrow{U}$ sont colinéaires, ce qui donne un sens à la notion de "position limite de la sécante".

PROPOSITION 20-1.9 La tangente à un arc paramétré $I \ni t \mapsto M(t)$ de classe C^k en un point t_0 où existe un premier indice fondamental p est la droite

$$T_{t_0} = M_0 + \text{vect}\left(\overrightarrow{M}^{(p)}(t_0)\right)$$

Cette droite est position limite de la sécante $M(t_0)M(t_0+h)$ lorsque h tend vers 0.

COROLLAIRE 20-1.10 Dans le cas de la dimension 2 et en un point t_0 régulier, si $M(t)$ a pour coordonnées $(x(t), y(t))$ dans le repère $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$, l'équation de la tangente en t_0 est donc, dans ce repère :

$$\begin{vmatrix} X - x(t_0) & x'(t_0) \\ Y - y(t_0) & y'(t_0) \end{vmatrix} = y'(t_0)(X - x(t_0)) - x'(t_0)(Y - y(t_0)) = 0$$

En un point où le premier indice fondamental est p, ce sont les dérivées $x^{(p)}(t_0)$ et $y^{(p)}(t_0)$ qui remplacent les dérivées premières.

REMARQUE 20-1.11 En un point où le premier indice fondamental n'existe pas, l'arc paramétré peut ne pas posséder de tangente, même s'il est de classe C^∞ :

Nous savons (*cf.* exercice **8-5**.20) que la fonction définie sur $\mathbb{R}^*$ par

$$f(t) = e^{-\frac{1}{t^2}}$$

et $f(0) = 0$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. L'arc paramétré $\mathcal{C}^\infty$ (dans $\mathbb{R}^2$) défini pour $t \in [-1, 1]$ par

$$M(t) = (f(t), 2f(t))$$

possède évidemment une tangente en $t = 0$ (son support est un segment dont $M(0)$ est une extrémité) bien que le premier indice fondamental ne soit pas défini en ce point. Par contre l'arc défini par

$$N(t) = (f(t), 0) \text{ pour } t \leqslant 0 \text{ et } N(t) = (0, f(t)) \text{ pour } t > 0$$

est également $\mathcal{C}^\infty$ et ne possède pas de tangente en $t = 0$ (son support est réunion de deux segments se rejoignant à l'origine, et on pourrait définir deux demi-tangentes en ce point).

20-1.2.3 Etude en un point birégulier : concavité, plan osculateur

Nous supposons dans cette section le point t_0 **birégulier**. Le développement limité en 0 à l'ordre 2 de l'application $h \mapsto \overrightarrow{M(t_0)M(t_0+h)} = \overrightarrow{M_0M_h}$ s'écrit alors

$$\overrightarrow{M_0M_h} = h\,\overrightarrow{M}'(t_0) + \frac{h^2}{2}\overrightarrow{M}''(t_0) + h^2\overrightarrow{\varepsilon}(h)$$

- Lorsque l'arc est plan, son support est inclus dans le plan

$$M_0 + \text{vect}\left(\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right)$$

et les coordonnées (X_h, Y_h) de M_h dans le repère local

$$\mathcal{R}_{t_0} = \left(M_0, \overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right)$$

vérifient

$$\begin{cases} X_h = h + o(h^2) \\ Y_h = \dfrac{h^2}{2} + o\left(h^2\right) \end{cases}$$

ce qui montre que, pour $h \neq 0$ petit, $Y_h > 0$. Le support de l'arc est donc localement inclus dans le demi-plan

$$M_0 + \mathbb{R}\,\overrightarrow{M}'(t_0) + \mathbb{R}^+\,\overrightarrow{M}''(t_0)$$

qu'on appelle **concavité** de l'arc au point t_0 (birégulier).

- Si $n = 3$, on appelle plan tangent à l'arc en t_0 tout plan contenant la tangente
$$T_{t_0} = M(t_0) + \mathbb{R}\overrightarrow{M}'(t_0)$$
Il y a donc une infinité de plans tangents. Parmi ceux-ci, le plan
$$M_0 + \operatorname{vect}\left(\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right)$$
est appelé **plan osculateur** en t_0. Le **demi-plan osculateur**
$$M_0 + \mathbb{R}\overrightarrow{M}'(t_0) + \mathbb{R}^+\overrightarrow{M}''(t_0)$$
est encore appelé **concavité de l'arc** en t_0. Il contient en effet (localement) la projection du support de l'arc sur le plan osculateur parallèlement à toute droite vectorielle non incluse dans $\operatorname{vect}\left(\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right)$: si on reprend le développement limité
$$\overrightarrow{M_0M_h} = h\,\overrightarrow{M}'(t_0) + \frac{h^2}{2}\overrightarrow{M}''(t_0) + h^2\overrightarrow{\varepsilon}(h)$$
et si $\overrightarrow{U} \notin \operatorname{vect}\left(\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right)$, les coordonnées de $h^2\overrightarrow{\varepsilon}(h)$ dans la base
$$\left(\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0), \overrightarrow{U}\right)$$
sont négligeables devant h^2, et les coordonnées de (X_h, Y_h, Z_h) de M_h dans le repère local $\left(M_0, \overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0), \overrightarrow{U}\right)$ vérifient
$$\begin{cases} X_h = h + o(h^2) \\ Y_h = \dfrac{h^2}{2} + o(h^2) \\ Z_h = o(h^2) \end{cases}$$
En projection sur le plan osculateur, on retrouve l'étude faite en dimension 2. Pour $|h|$ "petit", on observe une courbe qu'on peut confondre, en première approximation, avec la parabole d'équation $Y = \dfrac{X^2}{2}$.

20-1.2.4 Etude locale d'un arc plan

Pour un arc plan, en un point t_0 où existent les deux indices fondamentaux p et q, nous effectuons une étude dans le repère $\left(M_0, \overrightarrow{M}^{(p)}(t_0), \overrightarrow{M}^{(q)}(t_0)\right)$. Dans ce repère, la tangente à l'arc en t_0 est confondue avec l'axe des abscisses.

Le développement limité
$$\overrightarrow{M_0M_h} \underset{h\to 0}{=} \left(\frac{h^p}{p!} + o(h^p)\right)\overrightarrow{M}^{(p)}(t_0) + \left(\frac{h^q}{q!} + o(h^q)\right)\overrightarrow{M}^{(q)}(t_0)$$
montre que les coordonnées (X_h, Y_h) de M_h vérifient
$$X_h \sim \frac{h^p}{p!}$$
$$Y_h \sim \frac{h^q}{q!}$$
L'allure locale du support de l'arc dépend des parités de p et q :

- **Point ordinaire (à concavité) : p est impair et q est pair**

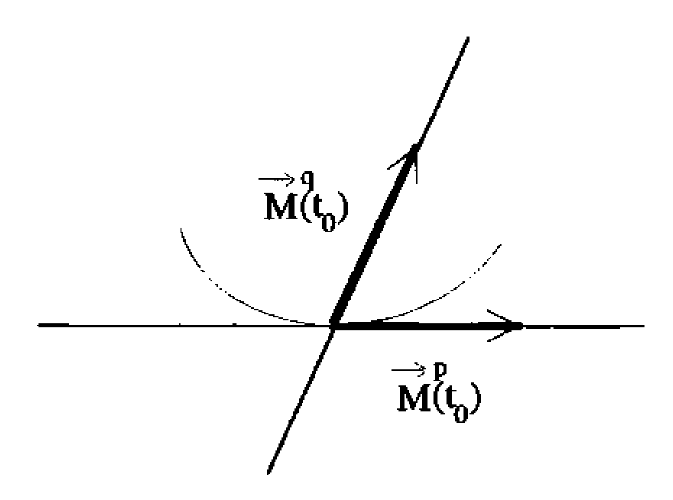


Figure **20.1** – *Point à concavité : p impair et q pair*

Le cas le plus fréquent est évidemment $p = 1$ et $q = 2$ (point birégulier). Lorsque h change de signe, il en est de même de X_h alors que Y_h garde un signe constant : le support de l'arc est situé localement dans le demi-plan $Y \geqslant 0$ (concavité de l'arc en t_0). La courbe reste du même côté de sa tangente.

Par contre, elle traverse toute autre droite (dite sécante) passant par M_0. En effet, dans le repère local une telle droite a une équation de la forme $\alpha X + \beta Y = 0$ avec $\alpha \neq 0$ et sépare le plan en deux demi-plans ouverts. Un point $M\,(X, Y)$ est dans l'un ou l'autre de ces demi-plans selon que la quantité $\alpha X + \beta Y$ est strictement positive ou négative. Pour le point M_h

$$\Delta\,(h) = \alpha X_h + \beta Y_h \sim \alpha \frac{h^p}{p!}$$

change de signe avec h, ce qui prouve que l'arc traverse la droite en M_0.

- **Point d'inflexion : p est impair et q est impair**

Dans la pratique, c'est le cas $p = 1$ et $q = 3$. Cette fois X_h et Y_h changent de signe avec h, et le support de l'arc traverse sa tangente en M_0. La courbe traverse en fait toute droite passant par M_0, puisque l'étude faite pour une sécante quelconque est la même que pour un point à concavité.

- **Point de rebroussement de première espèce : p est pair et q est impair**

Dans la pratique, nous aurons $p = 2$ et $q = 3$: le point de paramètre t_0 est stationnaire. Comme p est pair, X_h garde un signe constant et la courbe reste localement dans le demi-plan $X \geqslant 0$. Par contre, Y_h change de signe avec h, et le support de l'arc traverse la tangente. Pour une sécante quelconque d'équation $\alpha X + \beta Y = 0$ (avec $\alpha \neq 0$) on a, comme précédemment

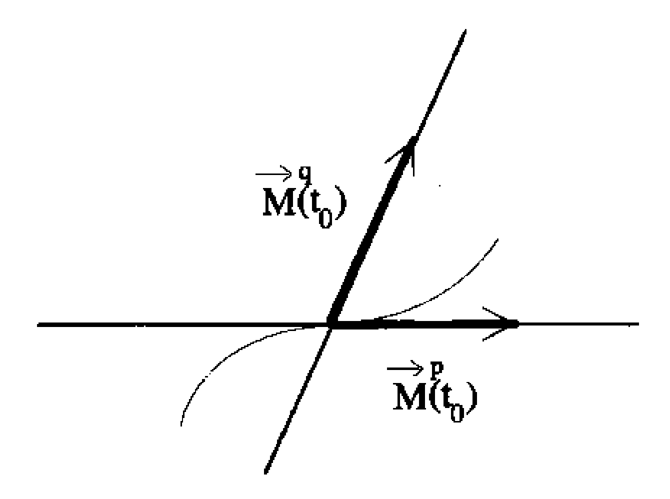


Figure **20.2** – *Point d'inflexion : p impair et q impair*

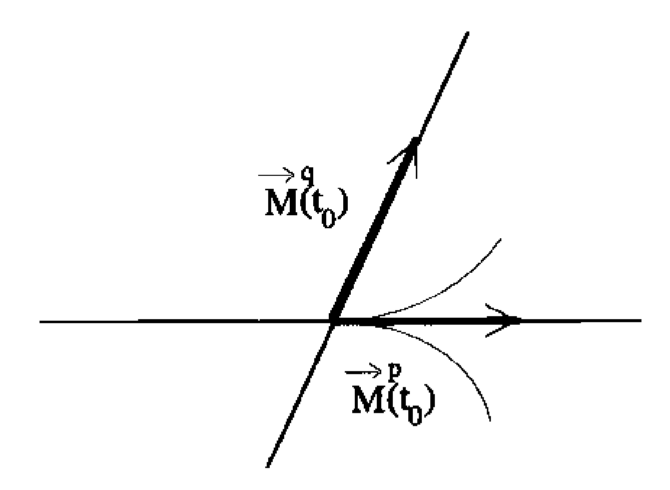


Figure **20.3** – *Point de rebroussement de première espèce : p pair et q impair*

$$\Delta(h) = \alpha X_h + \beta Y_h \sim \alpha \frac{h^p}{p!}$$

et cette quantité garde un signe constant : la courbe reste (localement) du même côté de chaque sécante (non tangente).

- **Point de rebroussement de seconde espèce : p est pair et q est pair**

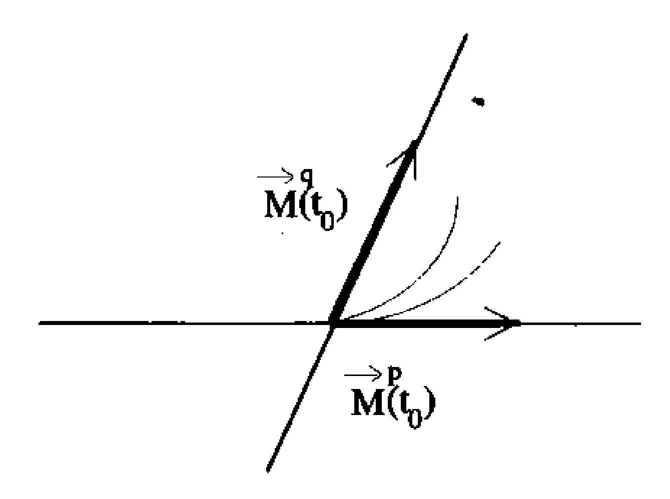


Figure **20.4** – *Point de rebroussement de seconde espèce : p pair et q pair*

Il s'agit le plus souvent d'un point stationnaire "exceptionnel", pour lequel $p = 2$ et $q = 4$. Comme p et q sont pairs, la courbe reste localement dans le quadrant $\{X \geqslant 0, Y \geqslant 0\}$, donc en particulier du même côté de la tangente. Pour une sécante quelconque d'équation $\alpha X + \beta Y = 0$ (avec $\alpha \neq 0$), nous avons

$$\Delta(h) = \alpha X_h + \beta Y_h \sim \alpha \frac{h^p}{p!}$$

Ici encore, cette quantité garde un signe constant, et la courbe reste (localement) du même côté de toute droite passant par M_0.

20-1.2.5 Conditions de contact

Nous nous plaçons ici dans le cas où $n = 2$ et nous considérons un arc paramétré

$$I \ni t \mapsto M(t) \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

Nous avons vu comment trouver l'équation de la tangente à un arc en un point régulier (ou plus généralement si le premier indice fondamental p existe) par un calcul de dérivée. Cette équation peut s'écrire

$$\begin{vmatrix} x - x(t_0) & x'(t_0) \\ y - y(t_0) & y'(t_0) \end{vmatrix} = y'(t_0)(x - x(t_0)) - x'(t_0)(y - y(t_0)) = 0$$

dans le cas où $p = 1$.

On peut parfois trouver cette équation en l'écrivant *a priori*, et en déterminant ses coefficients par une "condition de contact" :
Si $\mathcal{D}$ est une droite d'équation

$$(\mathcal{D}) : ax + by + c = 0 \quad (\text{avec } (a, b) \neq (0, 0))$$

considérons l'expression

$$\Delta(t) = a\,x(t) + b\,y(t) + c$$

$\Delta : I \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^k$. La droite $\mathcal{D}$ passe par $M(t_0)$ si et seulement si $\Delta(t_0) = 0$.

- Si $\Delta(t)$ est infiniment petit du premier ordre (par rapport à l'infiniment petit principal $t - t_0$) alors

  $$\Delta'(t_0) = a\,x'(t_0) + b\,y'(t_0) \neq 0$$

  et donc le vecteur $\overrightarrow{M}'(t_0)$ n'appartient pas à la direction de la droite $\mathcal{D}$. Celle-ci est donc sécante à l'arc au point de paramètre t_0.

- Si $\Delta(t)$ est infiniment petit négligeable par rapport à $t - t_0$, et si nous supposons t_0 régulier pour l'arc considéré, la condition $\Delta'(t_0) = 0$ montre que $\overrightarrow{M}'(t_0)$ dirige $\mathcal{D}$, qui est donc la tangente à l'arc au point de paramètre t_0. L'ordre de l'infiniment petit $\Delta(t)$ permet alors de préciser l'allure locale de l'arc au voisinage de t_0 :

  - Si $\Delta(t)$ est infiniment petit d'ordre pair (en général 2), il garde un signe constant au voisinage de t_0, et l'arc reste localement du même côté de sa tangente en t_0, qui est donc un point à concavité.
  - Si $\Delta(t)$ est infiniment petit d'ordre impair (en général 3), il change de signe avec $t - t_0$, et l'arc traverse sa tangente en t_0, qui est donc un point d'inflexion.

On utilise souvent ce genre de raisonnement lorsque $\Delta(t)$ est une fonction rationnelle de t. Dans ce cas, la notion d'ordre d'un infiniment petit correspond à la multiplicité d'une racine d'un polynôme :

EXEMPLE 20-1.12 Pour $\lambda \in \mathbb{R}$, on considère l'arc paramétré

$$(\mathcal{C}_\lambda) \quad \mathbb{R} - \{\lambda\} \ni t \mapsto M(t) \begin{pmatrix} x(t) = \dfrac{3t - t^3}{t - \lambda} \\ y(t) = \dfrac{1 - 3t^2}{t - \lambda} \end{pmatrix}$$

Montrer qu'en général, cet arc possède une unique tangente d'inflexion T_λ et trouver un arc régulier $\lambda \mapsto P(\lambda)$ dont la tangente au point de paramètre λ soit exactement T_λ (un tel arc est appelé **enveloppe des droites** T_λ).

- On pourrait ici déterminer **les points non bi-réguliers dits aussi points d'inflexion analytique** c'est à dire les valeurs du paramètre t vérifiant

$$\begin{vmatrix} x'(t) & x''(t) \\ y'(t) & y''(t) \end{vmatrix} = 0$$

Parmi les t vérifiant cette équation, on retrouve tous les points d'inflexion (pour lesquels on ne peut avoir $p = 1$ et $q = 2$, et donc pour lesquels le déterminant est certainement nul) mais aussi tous les points stationnaires (pour lesquels la première colonne du déterminant est nulle) ou encore des points pour lesquels les premiers indices fondamentaux seraient (par exemple) 1 et 4, qui sont des points à concavité.

- Il est préférable de chercher *a priori* l'équation d'une tangente en un point d'inflexion par la condition de contact :

Analyse : si $\mathcal{D}$ est une droite d'équation $ux + vy + w = 0$, et si $\mathcal{D}$ est tangente d'inflexion en un point de paramètre t_0, alors

$$\Delta(t) = u\frac{3t - t^3}{t - \lambda} + v\frac{1 - 3t^2}{t - \lambda} + w$$

est infiniment petit d'ordre $\geqslant 3$ au voisinage de t_0 (si c'était un infiniment petit d'ordre 1, $\mathcal{D}$ ne serait pas tangente en t_0 et s'il était infiniment petit d'ordre 2, l'arc resterait localement du même côté de $\mathcal{D}$, qui ne pourrait pas être tangente d'inflexion). Comme on suppose évidemment $t_0 \neq \lambda$, il est équivalent de dire que

$$(\lambda - t)\,\Delta(t) = u\,t^3 + 3v\,t^2 - (3u + w)\;t + (\lambda w - v)$$

est infiniment petit d'ordre $\geqslant 3$ en t_0. Comme $(u, v) \neq (0, 0)$, cette expression est un polynôme de degré 3 ou 2. On ne peut obtenir un infiniment petit d'ordre $\geqslant 3$ que si $u \neq 0$ et si ce polynôme possède t_0 comme racine triple, c'est à dire s'il s'écrit

$$u\,(t - t_0)^3 = u\left(t^3 - 3t_0\,t^2 + 3t_0^2\,t - t_0^3\right)$$

En identifiant les coefficients, et en prenant par exemple $u = 1$ (l'équation de $\mathcal{D}$ est définie à un scalaire multiplicatif non nul près), on obtient

$$\begin{cases} v = -t_0 \\ w + 3 = -3t_0^2 \\ v - \lambda w = t_0^3 \end{cases}$$

En reportant les valeurs de v et w dans la dernière équation, on obtient

$$\left(3t_0^2 + 3\right)\lambda = \left(t_0^3 + t_0\right)$$

et finalement

$$t_0 = 3\lambda$$

Synthèse : On prend $t_0 = 3\lambda$. Comme on doit avoir $t_0 \neq \lambda$, il faut choisir $\lambda \neq 0$. La droite T_λ a alors pour équation

$$x - 3\lambda\, y - 3\left(1 + 9\lambda^2\right) = 0$$

puisque $u = 1, v = -t_0$ et $w = -3\left(1 + t_0^2\right)$. Pour ce choix des paramètres, la quantité $\Delta(t)$ vaut

$$\Delta(t) = \frac{(t - 3\lambda)^3}{\lambda - t}$$

et est donc infiniment petit d'ordre 3 en t_0. De plus, on vérifie aisément que $y'(3\lambda) \neq 0$, ce qui montre que t_0 est un point régulier de l'arc. La droite T_λ est donc tangente à l'arc en t_0, et c'est une tangente d'inflexion puisque $\Delta(t)$ change de signe en t_0.

- **Enveloppe des droites $(T_\lambda)_{\lambda \neq 0}$:**

Analyse : Supposons qu'un arc Γ de classe $\mathcal{C}^1$

$$\mathbb{R}^* \ni \lambda \mapsto P(\lambda) \begin{pmatrix} X(\lambda) \\ Y(\lambda) \end{pmatrix}$$

soit tel que, pour tout λ, la tangente en λ à Γ soit la droite T_λ. Comme $P(\lambda)$ doit appartenir à cette droite, on a nécessairement

$$\forall \lambda \in \mathbb{R}^* \quad X(\lambda) - 3\lambda\, Y(\lambda) - 3\left(1 + 9\lambda^2\right) = 0 \qquad (*)$$

De plus, pour tout λ, le vecteur $\overrightarrow{P}'(\lambda)$ doit appartenir à la direction de la droite T_λ, ce qui se traduit par

$$\forall \lambda \in \mathbb{R}^* \quad X'(\lambda) - 3\lambda\, Y'(\lambda) = 0$$

En dérivant la première identité par rapport à λ et en comparant à la seconde, on obtient

$$\forall \lambda \in \mathbb{R}^* \quad -3\,Y(\lambda) - 3\,(18\lambda) = 0 \qquad (**)$$

ce qui donne $Y(\lambda) = -18\lambda$, et en reportant dans la première identité

$$\forall \lambda \in \mathbb{R}^* \quad X(\lambda) = 3\lambda\, Y(\lambda) + 3\left(1 + 9\lambda^2\right) = 3 - 27\lambda^2$$

Synthèse : L'arc défini par

$$\mathbf{R}^* \ni \lambda \mapsto P(\lambda) \begin{pmatrix} 3 - 27\lambda^2 \\ -18\lambda \end{pmatrix}$$

est évidemment régulier. Son support est une parabole privée de son sommet, puisque

$$X(\lambda) = 3 - \frac{1}{12} Y(\lambda)^2$$

Il répond évidemment à la question puisque, vérifiant les identités $(*)$ et $(**)$, il vérifie

$$\forall \lambda \in \mathbb{R}^* \quad X'(\lambda) - 3\lambda\, Y'(\lambda) = 0$$

identité qu'on obtient ici encore en dérivant $(*)$ et en comparant avec $(**)$.

REMARQUE 20-1.13 La démarche suivie dans l'exemple précédent est générale pour résoudre les problèmes d'enveloppe : soit I est un intervalle de $\mathbb{R}$ et

$$(\mathcal{D}_t)_{t\in I} : \quad u(t)\, x + v(t)\, y + w(t) = 0$$

une famille de droites dépendant d'un paramètre t avec u, v et $w \in C^1(I, \mathbb{R})$. Si un arc de classe C^1

$$I \ni t \mapsto M(t) \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

est tel que pour tout $t \in I$ la tangente en t est $\mathcal{D}_t$, on aura nécessairement

$$\forall t \in I \quad \begin{cases} u(t)\, x(t) + v(t)\, y(t) + w(t) = 0 \\ u(t)\, x'(t) + v(t)\, y'(t) = 0 \end{cases}$$

En dérivant la première identité et en comparant à la seconde, on obtient

$$\forall t \in I \quad \begin{cases} u(t)\, x(t) + v(t)\, y(t) + w(t) = 0 \\ u'(t)\, x(t) + v'(t)\, y(t) + w'(t) = 0 \end{cases}$$

Lorsque $(u'(t), v'(t))$ ne s'annule pas, le point $M(t)$ doit donc être cherché à l'intersection de la droite $\mathcal{D}_t$ et de la droite $\mathcal{D}'_t$ d'équation

$$(\mathcal{D}'_t) : \quad u'(t)\, x + v'(t)\, y + w'(t) = 0$$

Cette intersection est réduite à un point pour tout $t \in I$ si et seulement si

$$\forall t \in I \quad \begin{vmatrix} u(t) & v(t) \\ u'(t) & v'(t) \end{vmatrix} \neq 0$$

On fait ensuite une synthèse comme dans l'étude précédente.

REMARQUE 20-1.14 En dimension 3, on peut également étudier le contact entre un arc paramétré et un plan donné par une équation

$$(\mathcal{P}) : \quad ax + by + cz + d = 0$$

Au voisinage d'un point t_0 trirégulier d'un arc de classe C^k $(k \geqslant 3)$

$$I \ni t \mapsto M(t) \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

on étudie la quantité

$$\Delta(t) = a\, x(t) + b\, y(t) + c\, z(t) + d$$

Si $\Delta(t)$ est un infiniment petit d'ordre 1 (par rapport à $t - t_0$), le plan $\mathcal{P}$ passe par $M(t_0)$ mais n'est pas tangent à l'arc en ce point. Si $\Delta(t)$ est infiniment petit d'ordre 2, le plan $\mathcal{P}$ est tangent à l'arc en t_0, et c'est le plan osculateur en t_0 si cet infiniment petit est d'ordre 3.

EXERCICE 20-1.15 On donne l'arc γ de $\mathbb{R}^3$ défini par la paramétrisation

$$x(t) = \frac{1}{1+t^2} \quad y(t) = \frac{t}{1+t^2} \quad z(t) = \frac{t^3}{1+t^2}$$

Montrer que cet arc est simple et régulier. Montrer que par un point M de l'espace, il passe en général 1, 2 ou 3 plans osculateurs à γ. Lorsqu'il y en a trois, correspondant aux points $(M_i)_{1\leqslant i\leqslant 3}$ de paramètres $(t_i)_{1\leqslant i\leqslant 3}$, montrer que A, M_1, M_2 et M_3 sont coplanaires.

20-1.3 Branches infinies, asymptotes

Nous nous limiterons ici à la dimension 2.

DÉFINITION 20-1.16 *Soit* $\Gamma : I \ni t \mapsto M(t)$ *un arc paramétré continu et* $t_0 \in \overline{I} - I$ *un point adhérent à* I *(avec la possibilité* $t_0 = +\infty$ *si* I *n'est pas majoré,* $t_0 = -\infty$ *si* I *est non minoré). On dit que* Γ *présente une branche infinie au voisinage de* t_0 *si et seulement si*

$$\lim_{t \to t_0} \left\| \overrightarrow{OM(t)} \right\| = +\infty$$

Il est clair que ceci ne dépend pas de l'origine O choisie dans $\mathcal{E}_2$.

La notion de direction asymptotique correspond à la situation où, lorsque t tend vers t_0, le point $M(t)$ s'éloigne à l'infini "dans une direction donnée" :

DÉFINITION 20-1.17 *Soit* $\Gamma : I \ni t \mapsto M(t)$ *un arc paramétré continu présentant une branche infinie au voisinage de* t_0. *Lorsque la limite*

$$\lim_{t \to t_0} \frac{\overrightarrow{OM(t)}}{\left\| \overrightarrow{OM(t)} \right\|} = \overrightarrow{U}$$

existe (on a alors évidemment $\left\| \overrightarrow{U} \right\| = 1$*), elle ne dépend alors pas de l'origine choisie. La droite* $OM(t)$ *possède une position limite pour* $t \to t_0$ *(voir exercice* **20-1**.*8) qui est la droite* $O + \text{vect}\left(\overrightarrow{U}\right)$. *On dit alors que l'arc* Γ *possède la direction asymptotique* $\text{vect}\left(\overrightarrow{U}\right)$ *en* t_0.

Démonstration : Si la limite existe et si O_1 est un point quelconque de $\mathcal{E}_2$, on a

$$\left\| \overrightarrow{OM(t)} \right\| - \left\| \overrightarrow{OO_1} \right\| \leqslant \left\| \overrightarrow{O_1M(t)} \right\| \leqslant \left\| \overrightarrow{OM(t)} \right\| + \left\| \overrightarrow{OO_1} \right\|$$

ce qui donne évidemment

$$\left\| \overrightarrow{O_1M(t)} \right\| \underset{t \to t_0}{\sim} \left\| \overrightarrow{OM(t)} \right\|$$

et il suffit d'écrire

$$\frac{\overrightarrow{O_1M(t)}}{\left\| \overrightarrow{O_1M(t)} \right\|} = \frac{\left\| \overrightarrow{OM(t)} \right\|}{\left\| \overrightarrow{O_1M(t)} \right\|} \frac{\overrightarrow{OM(t)}}{\left\| \overrightarrow{OM(t)} \right\|} + \frac{\overrightarrow{O_1O}}{\left\| \overrightarrow{O_1M(t)} \right\|}$$

pour voir que

$$\lim_{t \to t_0} \frac{\overrightarrow{O_1M(t)}}{\left\| \overrightarrow{O_1M(t)} \right\|} = \lim_{t \to t_0} \frac{\overrightarrow{OM(t)}}{\left\| \overrightarrow{OM(t)} \right\|} = \overrightarrow{U}$$

et le choix de l'origine O n'importe donc pas dans la définition précédente. ■

Lorsque l'on travaille dans un repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$ avec $M(t)$ de coordonnées $(x(t), y(t))$

$$\overrightarrow{OM(t)} = x(t)\ \overrightarrow{i} + y(t)\ \overrightarrow{j}$$

l'arc présente une branche infinie en t_0 si et seulement si

$$\lim_{t \to t_0} |x(t)| + |y(t)| = +\infty$$

Si $t \mapsto x(t)$ ne s'annule pas au voisinage de t_0, on peut repérer la droite $OM(t)$ par son coefficient directeur

$$m(t) = \frac{y(t)}{x(t)}$$

- Si on a

  $$\lim_{t_0} m(t) = \pm\infty$$

  alors la position limite de $OM(t)$ est l'axe $Oy = O + \text{vect}\left(\overrightarrow{j}\right)$, et la direction asymptotique est donc vect $\left(\overrightarrow{j}\right)$. On a bien $x(t) \underset{t_0}{=} o(y(t))$ dans ce cas, donc $\overrightarrow{OM(t)} \underset{t_0}{\sim} y(t)\ \overrightarrow{j}$ et par conséquent

  $$\lim_{t_0} \frac{\overrightarrow{OM(t)}}{\left\|\overrightarrow{OM(t)}\right\|} = \pm\overrightarrow{j}$$

  le signe dépendant de celui de $y(t)$ au voisinage de t_0.

- De même, si

  $$\lim_{t_0} m(t) = \alpha \in \mathbb{R}$$

  la droite $OM(t)$ a pour position limite la droite d'équation $y = \alpha x$. On a en effet dans ce cas

  $$\overrightarrow{OM(t)} \underset{t_0}{\sim} x(t) \left(\overrightarrow{i} + \alpha\, \overrightarrow{j}\right)$$

  et donc

  $$\lim_{t_0} \frac{\overrightarrow{OM(t)}}{\left\|\overrightarrow{OM(t)}\right\|} = \pm \frac{\overrightarrow{i} + \alpha\, \overrightarrow{j}}{\left\|\overrightarrow{i} + \alpha\, \overrightarrow{j}\right\|}$$

  (le signe dépend de celui de $x(t)$ au voisinage de t_0) qui est bien un vecteur directeur de la droite d'équation $y = \alpha x$.

Lorsque l'arc présente une direction asymptotique vect $\left(\overrightarrow{U}\right)$ en t_0, **on cherche s'il existe une droite (dirigée par $\overrightarrow{U}$) qui est asymptote à l'arc :**

- Si la direction asymptotique est vect $\left(\overrightarrow{J}\right)$ et si

$$\lim_{t\to t_0} x(t) = x_0$$

on dit que la droite d'équation $x = x_0$ est asymptote ("verticale") à l'arc. Si par contre $\lim\limits_{t\to t_0} x(t) = \pm\infty$, on dit parfois que l'arc présente une branche parabolique dans la direction de l'axe Oy.

- Si la direction asymptotique est celle de la droite d'équation $y = \alpha x$, et si la limite

$$\lim_{t\to t_0} y(t) - \alpha\, x(t) = \beta$$

existe (dans $\mathbb{R}$), on dit que la droite Δ d'équation

$$(\Delta): \quad y = \alpha x + \beta$$

est asymptote à l'arc en t_0. Si cette limite est $\pm\infty$, on parlera de branche parabolique dans la direction de la droite d'équation $y = \alpha\, x$.

EXERCICE 20-1.18 Plus précisément, si un arc continu $t \mapsto M(t)$ présente une branche infinie en t_0, on dit qu'une droite $\mathcal{D}$ est asymptote à l'arc en t_0 si et seulement si

$$\lim_{t\to t_0} d(M(t), \mathcal{D}) = 0$$

(où $d(M(t), \mathcal{D})$ est la distance du point $M(t)$ à la droite $\mathcal{D}$). Montrer que cette propriété ne dépend pas de la norme choisie sur l'espace vectoriel sous-jacent à $\mathcal{E}_2$. Montrer que, si une asymptote existe, elle est unique, et que l'arc présente alors en t_0 une direction asymptotique qui est celle de $\mathcal{D}$. Montrer également qu'alors, en choisissant un repère quelconque du plan, on se retrouve dans une des deux situations décrites précédemment.

DANS LA PRATIQUE :

On essaie souvent d'obtenir des développements limités (ou asymptotiques) de $x(t_0+h)$ et $y(t_0+h)$ pour h tendant vers 0 (si $t_0 = \pm\infty$, on utilisera un infiniment petit qui sera souvent $h = \dfrac{1}{t}$ mais peut, selon les cas, être un autre infiniment petit de référence, comme $\dfrac{1}{\ln|t|}$ ou $e^{-|t|}$). On peut avoir une asymptote "oblique" lorsque ces deux quantités sont des infiniment grands simultanés. Le premier terme de chacun des ces développements donne un équivalent simple de $x(t_0+h)$ et $y(t_0+h)$ au voisinage de 0, et permet de voir si ces infiniment grands sont de même ordre. Si c'est le cas[3], on trouve aisément $\alpha \in \mathbb{R}^*$ avec

$$y(t_0+h) \underset{h\to 0}{\sim} \alpha\, x(t_0+h)$$

ce qui détermine une direction asymptotique. Lorsque la différence admet un développement de la forme

$$y(t_0+h) - \alpha\, x(t_0+h) = \beta + \gamma h^p + o(h^p)$$

[3]Dans le cas contraire, l'arc présente une branche parabolique dans la direction de Oy ou de Ox, selon que $y(t) \underset{t_0}{=} o(x(t))$ ou $x(t) \underset{t_0}{=} o(y(t))$.

avec $\beta \in \mathbb{R}, p \in \mathbb{N}^*$ et $\gamma \neq 0$, l'arc possède une asymptote[4] d'équation $y = \alpha x + \beta$. **Le signe de γ ainsi que la parité de p permettent de placer l'arc** (au voisinage de t_0) **par rapport à son asymptote** :

$$\delta(h) = y(t_0 + h) - \alpha x(t_0 + h) - \beta \underset{h \to 0}{\sim} \gamma h^p$$

et l'équivalent donne le signe (au voisinage de 0). Lorsque $\delta(h) > 0$, la courbe est située "au dessus" de son asymptote. Elle est en dessous lorsque $\delta(h) < 0$. On remarquera que, lorsque p est pair, $\delta(h)$ ne change pas de signe avec h.

EXEMPLE 20-1.19 Etudier la branche infinie, pour $t \to 1$, pour l'arc paramétré $t \mapsto M(t)$ défini par

$$x(t) = \frac{t-2}{t(t-1)} \quad \text{et} \quad y(t) = \frac{(2t-1)t}{(t-2)(t-1)}$$

Au voisinage de 1, on a évidemment

$$x(t) \sim -\frac{1}{(t-1)} \quad \text{et} \quad y(t) \sim -\frac{1}{t-1}$$

ce qui montre l'existence d'une direction asymptotique qui est celle de la première bissectrice. En posant $t = 1 + h$, on obtient

$$\begin{cases} x(t) = -\dfrac{1}{h}\left(\dfrac{1-h}{1+h}\right) = -\dfrac{1}{h}\left(1 - 2h + 2h^2 + o(h^2)\right) \\ y(t) = -\dfrac{1}{h}\left(\dfrac{(1+2h)(1+h)}{(1-h)}\right) = -\dfrac{1}{h}\left(1 + 4h + 6h^2 + o(h^2)\right) \end{cases}$$

ce qui donne

$$y(t) - x(t) = -6 - 4h + o(h)$$

L'arc considéré présente donc une asymptote d'équation

$$y = x - 6$$

Pour t tendant vers 1 par valeurs supérieures, $x(t)$ et $y(t)$ tendent vers $-\infty$ et l'arc est situé sous son asymptote[5]. Il est situé au dessus pour t tendant vers 1 par valeurs inférieures, avec $x(t)$ et $y(t)$ qui tendent vers $+\infty$.

[4]Lorsque $y(t_0 + h) - \alpha x(t_0 + h)$ est encore un infiniment grand pour $h \to 0$, l'arc présente une branche parabolique dans la direction de la droite d'équation $y = \alpha x$.

[5]On a là une information locale. Comme la quantité

$$\Delta(t) = y(t) - x(t) + 6$$

est rationnelle, on pourrait faire une étude de signe de cette différence pour voir que, **globalement**,

$$\Delta(t) > 0 \text{ pour } t \in]-\infty, 0[\cup \left]\frac{1}{2}, 1\right[\cup]2, +\infty[$$

et que la courbe traverse l'asympote pour $t = \dfrac{1}{2}$.

20-1.4 Plan d'étude d'un arc en dimension 2

- Pour permettre un tracé sommaire du support d'un arc paramétré ou contrôler ce qui est affiché par l'écran d'un ordinateur ou d'une calculatrice, on commence par déterminer **le domaine de définition** $\mathcal{D}$ et la régularité de l'application $t \mapsto M(t)$.

- On **restreint ensuite éventuellement le domaine d'étude** en tenant compte des propriétés de symétrie (parité, imparité, périodicité simultanément pour les fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$). Par exemple, si $t \mapsto x(t)$ est paire et $t \mapsto y(t)$ est impaire, on fait l'étude sur $\mathcal{D} \cap \mathbb{R}^+$, et on complète le tracé du support par symétrie par rapport à l'axe des abscisses.

- On effectue alors **l'étude des variations des fonctions** $t \mapsto x(t)$ et $t \mapsto y(t)$ (en général en utilisant le calcul des dérivées). On fait ainsi apparaître les points remarquables (à tangentes verticales ou horizontales) ainsi que les points stationnaires ($x'(t) = y'(t) = 0$).

- **Etude des des points stationnaires** : si t_0 est un point stationnaire, c'est en général un développement limité au voisinage de 0 des fonctions $h \mapsto x(t_0 + h)$ et $h \mapsto y(t_0 + h)$ qui permet d'obtenir (lorsqu'ils existent) les deux indices fondamentaux p et q. On peut alors placer la tangente et avoir l'allure locale du support de l'arc, conformément aux résultats de la section **20-1.2.4**.

- **Etude "aux bornes"** de l'intervalle : ce sont souvent des branches infinies, mais on rencontre aussi parfois des "points d'arrêts" : lorsque $t \to t_0$ (avec $t_0 \in \overline{I} - I$), on a

  $$(x(t), y(t)) \to (x_0, y_0)$$

  ce qui permet de considérer le point $M_0(x_0, y_0)$ comme un point limite appartenant au support de l'arc. On pourra placer une tangente éventuelle en ce point en étudiant la position limite d'une sécante, c'est à dire en déterminant

  $$\lim_{t \to t_0} \frac{y(t) - y_0}{x(t) - x_0}$$

  Si l'étude fait apparaître des asymptotes, il est important de placer (en général localement) l'arc par rapport à ces asymptotes.

- Lorsque le tracé sommaire fait apparaître des **points d'inflexion ou multiples** (où se croisent deux branches du support de l'arc), on peut terminer l'étude (si les calculs ne sont pas trop compliqués) en déterminant ces points par les méthodes qui suivent :

  - Nous avons vu comment les conditions de contact permettent parfois de déterminer les points d'inflexion. On peut aussi les chercher parmi les points non biréguliers (dits d'inflexion "analytique"), vérifiant

    $$\begin{vmatrix} x'(t) & x''(t) \\ y'(t) & y''(t) \end{vmatrix} = 0$$

Ceci peut s'écrire aussi, si $x'(t) \neq 0$

$$m'(t) = 0 \quad \text{avec} \quad m(t) = \frac{y'(t)}{x'(t)}$$

Le réel $m(t)$ représente le coefficient directeur de la tangente en t. Lorsque $m'(t)$ s'annule en t_0 en changeant de signe ($m(t)$ passe alors par un extremum), l'arc présente effectivement un point d'inflexion en t_0.

En effet, pour h non nul suffisamment petit, la quantité

$$\delta(h) = (y(t_0 + h) - y(t_0)) - m(t_0)(x(t_0 + h) - x(t_0))$$

(dont le signe permet de placer le point $M(t_0 + h)$ par rapport à la tangente en t_0) s'écrit

$$\begin{aligned}\delta(h) &= (x(t_0 + h) - x(t_0)) \left[\frac{(y(t_0 + h) - y(t_0))}{(x(t_0 + h) - x(t_0))} - m(t_0)\right] \\ &\sim hx'(t_0) \left[\frac{(y(t_0 + h) - y(t_0))}{(x(t_0 + h) - x(t_0))} - m(t_0)\right]\end{aligned}$$

(à cause de l'hypothèse $x'(t_0) \neq 0$). D'après la formule des accroissements finis généralisée (voir théorème **8-5**.6) on peut écrire

$$\left[\frac{(y(t_0 + h) - y(t_0))}{(x(t_0 + h) - x(t_0))} - m(t_0)\right] = m(t_h) - m(t_0)$$

avec $t_h \in]t_0, t_0 + h[$ ou $]t_0 + h, t_0[$ (selon le signe de h). Si $m(t)$ est extrémal en t_0 alors $m(t_h) - m(t_0)$ a un signe constant, et $\delta(h)$ change bien de signe avec h.

Par contre, lorsque $m'(t)$ s'annule en t_0 sans changer de signe, le raisonnement qui précède montre que le point t_0 n'est pas point d'inflexion. On peut retenir :

En un point d'un arc régulier de classe C^2 où le coefficient directeur de la tangente est défini et passe par un extremum, l'arc présente un point d'inflexion.

– Les points multiples se déterminent en résolvant le système

$$\begin{cases} x(t) = x(t') \\ y(t) = y(t') \end{cases} \quad \text{avec} \quad t, t' \in I \text{ et } t \neq t'$$

Nous verrons sur l'exemple qui suit que, lorsque $x(t)$ et $y(t)$ sont des fractions rationnelles en t, on cherche à déterminer t et t' par leur somme et leur produit.

EXERCICE 20-1.20 Etudier l'arc paramétré défini par

$$x(t) = \frac{t(t+2)}{2t+2} \quad \text{et} \quad y(t) = \frac{t+2}{t(t+1)}$$

On obtient les résultats suivants :

- L'arc est de classe C^∞ sur son domaine de définition $\mathbb{R}-\{0,-1\}$. Il n'y a pas de symétrie apparente permettant de restreindre l'ensemble d'étude.

- $x'(t) = \frac{1}{2}\frac{t^2+2t+2}{(t+1)^2}$ et $y'(t) = -\frac{t^2+4t+2}{t^2(t+1)^2}$. L'arc est régulier, la fonction $t \mapsto x(t)$ est strictement croissante sur chacun des intervalles de son domaine de définition. La fonction $y'(t)$ s'annule pour $t = -2 \pm \sqrt{2}$ et $t \mapsto y(t)$ est strictement croissante sur $\left]-2-\sqrt{2},-1\right[$ et $\left]-1,-2+\sqrt{2}\right[$, strictement décroissante sur $\left]-\infty,-2-\sqrt{2}\right[$, sur $\left]-2+\sqrt{2},0\right[$ et sur $]0,+\infty[$. (Faire un tableau de variations regroupant tous ces résultats, ce qui est commode ensuite pour le tracé du support).

- **Branches infinies :**

  Pour $t \to \pm\infty$ on a $x(t) \to \pm\infty$ et $y(t) \to 0$, avec le même signe que t. On a donc une asymptote horizontale d'équation $y = 0$, la courbe étant sous l'asymptote pour $t \to -\infty$, au dessus pour $t \to +\infty$. On remarque que la courbe traverse cette asymptote pour $t = -2$.

  Pour $t \to -1$, on pose $t + 1 = h$, et on a

$$x(-1+h) = -\frac{1}{2h} + \frac{h}{2} \quad \text{et} \quad y(-1+h) = \frac{1+h}{(-1+h)h} = -\frac{1}{h} - 2 - 2h + o(h)$$

  ce qui donne d'abord

$$y(-1+h) \sim 2x(-1+h)$$

  puis

$$y(-1+h) - 2x(-1+h) = -2 - 3h + o(h)$$

  On obtient donc une asymptote oblique d'équation $y = 2x - 2$, la courbe étant sous l'asymptote lorsque $t \to 1$ par valeurs supérieures, au dessus pour $t \to 1$ par valeurs inférieures. En formant

$$y(t) - 2x(t) + 2 = \frac{t+2}{t(t+1)} - 2\frac{t(t+2)}{2t+2} + 2 = -\frac{t^2-t-2}{t}$$

  on s'aperçoit que l'arc traverse son asymptote pour $t = 2$.

- Un tracé du support laisse apparaître un point d'inflexion pour une valeur de $t \in]-\infty,-2[$. On peut vérifier ce résultat en calculant le coefficint directeur de la tangente

$$m(t) = \frac{y'(t)}{x'(t)} = -2\frac{t^2+4t+2}{t^2(t^2+2t+2)} \text{ et } m'(t) = 4\frac{(t+1)\left(t^3+6t^2+6t+4\right)}{t^3(t^2+2t+2)^2}$$

  (il est normal de trouver $t+1$ en facteur, puisque pour $t = -1$ l'asymptote, qui joue le rôle de tangente à l'infini, est traversée par l'arc). Une étude des variations du polynôme t^3+6t^2+6t+4 montre qu'il possède une unique racine réelle $t \approx -4,9514$. En ce point, $m'(t)$ s'annule bien en changeant de signe.

- Enfin, le tracé montre l'existence d'un point double. On résout donc

$$\frac{t(t+2)}{2t+2} = \frac{t'(t'+2)}{2t'+2} \quad \text{et} \quad \frac{t+2}{t(t+1)} = \frac{t'+2}{t'(t'+1)} \text{ avec } t \neq t'$$

Comme ces égalités donnent des relations de proportionnalité entre numérateurs et dénominateurs, nous aurons aussi

$$x(t) = x(t') = \frac{t(t+2) - t'(t'+2)}{(2t+2) - (2t'+2)} = \frac{(t+t')}{2} + 1 \text{ et}$$

$$x(t) = x(t') = \frac{t(t+2) + t'(t'+2)}{(2t+2) + (2t'+2)} = \frac{(t+t')^2 + 2(t+t') - 2tt'}{2(t+t') + 4}$$

La même manipulation donne

$$y(t) = y(t') = \frac{(t+2) - (t'+2)}{t(t+1) - t'(t'+1)} = \frac{1}{t+t'+1} \text{ et}$$

$$y(t) = y(t') = \frac{t+2+t'+2}{t(t+1) + t'(t'+1)} = \frac{(t+t')+4}{(t+t')^2 + (t+t') - 2tt'}$$

On détermine alors t et t' par leur somme S et leur produit P qui vérifient

$$\frac{S}{2} + 1 = \frac{S^2 + 2S - 2P}{2S+4} \text{ et } \frac{1}{S+1} = \frac{S+4}{S^2 + S - 2P}$$

On obtient aisément

$$2S + 2P + 4 = 0 \text{ et } 4S + 2P + 4 = 0$$

soit $S = 0$ et $P = -2$. On obtient donc $\{t, t'\} = \{\pm\sqrt{2}\}$, et les coordonnées du point double sont

$$\left(\frac{S}{2} + 1, \frac{1}{S+1}\right) = (1, 1)$$

EXERCICE 20-1.21 Soit $a > 0$. Etudier l'arc du plan euclidien défini dans un repère orthonormé par

$$x(t) = 4a\cos^3 t \quad \text{et} \quad y(t) = 4a\sin^3 t$$

Etudier notamment les points singuliers. Montrer que la portion de tangente comprise entre l'axe des abscisses et l'axe des ordonnées est de longueur constante. En remarquant que

$$x(t) + i\,y(t) = 3a\,e^{it} + a\,e^{-3it}$$

montrer que le support de cet arc est trajectoire d'un point fixe d'un cercle de rayon a qui roule sans glisser à l'intérieur d'un cercle de rayon $4a$ (hypocycloïde à 4 rebroussements ou astroïde).

20-1.5 Cas des arcs plans définis en coordonnées polaires

20-1.5.1 Rappel : coordonnées polaires

Le plan euclidien $\mathcal{E}_2$ orienté est rapporté à un repère orthonormé direct $\left(O, \vec{\imath}, \vec{\jmath}\right)$. La droite $O + \mathbb{R}\vec{\imath}$ est appelée axe polaire. Si $\theta \in \mathbb{R}$, on note $\vec{u}(\theta)$ le vecteur unitaire image de $\vec{\imath}$ par la rotation d'angle θ

$$\vec{u}(\theta) = \cos\theta\,\vec{\imath} + \sin\theta\,\vec{\jmath}$$

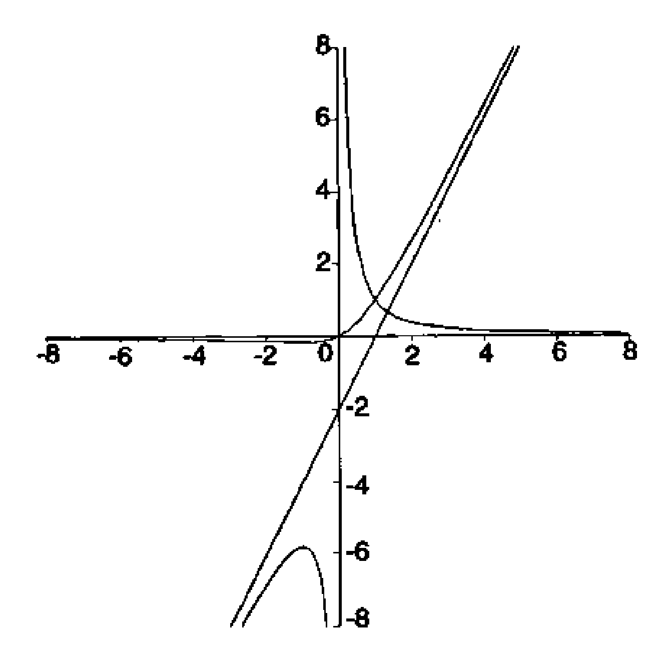


Figure **20.5** – *Arc paramétré* $x(t) = \dfrac{t(t+2)}{2t+2}$ *et* $y(t) = \dfrac{t+2}{t(t+1)}$

et $\overrightarrow{v}(\theta) = \overrightarrow{u}\left(\theta + \frac{\pi}{2}\right)$ le vecteur s'en déduisant par rotation de $\frac{\pi}{2}$, c'est à dire tel que la famille $(\overrightarrow{u}(\theta), \overrightarrow{v}(\theta))$ soit une base orthonormale directe du plan vectoriel sous-jacent

$$\overrightarrow{v}(\theta) = -\sin\theta\,\overrightarrow{I} + \cos\theta\,\overrightarrow{J}$$

Les applications $\theta \mapsto \overrightarrow{u}(\theta)$ et $\theta \mapsto \overrightarrow{v}(\theta)$ sont évidemment de classe C^∞ sur $\mathbb{R}$ et vérifient

$$\frac{d\overrightarrow{u}}{d\theta} = \overrightarrow{v}(\theta) \quad \text{et} \quad \frac{d\overrightarrow{v}}{d\theta} = -\overrightarrow{u}(\theta)$$

DÉFINITION 20-1.22 *Soit $M \in \mathcal{E}_2$. Un couple $(r, \theta) \in \mathbb{R}^2$ est un système de coordonnées polaires de M par rapport au repère $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$ si et seulement si*

$$\overrightarrow{OM} = r\,\overrightarrow{u}(\theta)$$

Si (x, y) sont les coordonnées de M dans $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$, ceci équivaut évidemment à

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

Si $M = O$, on voit que les systèmes de coordonnées polaires qui conviennent sont les couples $(0, \theta)$, avec $\theta \in \mathbb{R}$ quelconque. Si $M \neq O$, et si (r, θ) est un système de coordonnées polaires de M, on notera que l'égalité $\overrightarrow{OM} = r\,\overrightarrow{u}(\theta)$ définit r comme une mesure algébrique, avec $\left\|\overrightarrow{OM}\right\| = |r|$. On a donc deux possibilités :

- $r > 0$ et donc $r = \left\|\overrightarrow{OM}\right\| = \sqrt{x^2+y^2}$ avec

$$\overrightarrow{u}(\theta) = \frac{\overrightarrow{OM}}{\left\|\overrightarrow{OM}\right\|} = \frac{x}{\sqrt{x^2+y^2}}\overrightarrow{\imath} + \frac{y}{\sqrt{x^2+y^2}}\overrightarrow{\jmath}$$

Le réel θ est alors unique modulo 2π : l'égalité précédente signifie simplement que θ est une détermination de l'argument du nombre complexe de module 1

$$z = \frac{x}{\sqrt{x^2+y^2}} + i\frac{y}{\sqrt{x^2+y^2}}$$

- $r < 0$ et donc $r = -\left\|\overrightarrow{OM}\right\| = -\sqrt{x^2+y^2}$. Avec les notations précédentes, on doit avoir

$$e^{i\theta} = -z = -\left(\frac{x}{\sqrt{x^2+y^2}} + i\frac{y}{\sqrt{x^2+y^2}}\right)$$

ce qui montre que, cette fois, $\theta + \pi$ doit être une détermination de l'argument de z.

On retiendra :

PROPOSITION 20-1.23 Si M est un point du plan distinct de l'origine, M admet une infinité de systèmes de coordonnées polaires dans le repère $\left(O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right)$. Si (r,θ) est un tel système, les autres sont de la forme

$$(r, \theta + 2k\pi) \quad \textbf{ou} \quad (-r, \theta + \pi + 2k\pi) \quad \textbf{(avec } k \in \mathbb{Z}\textbf{)}$$

20-1.5.2 Paramétrage polaire d'un arc

Nous cherchons des conditions suffisantes pour pouvoir représenter simplement un arc en utilisant les coordonnées polaires. Dans tout ce qui suit, le plan $\mathcal{E}_2$ est rapporté à un repère orthonormé direct $\left(O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right)$.

PROPOSITION 20-1.24 Soit $I \ni t \mapsto M(t)$ un arc de classe C^k (avec $k \geqslant 1$), dont le support ne contient pas l'origine O. Il existe un couple de fonctions numériques de classe C^k

$$t \mapsto \rho(t) \quad \textbf{et} \quad t \mapsto \psi(t)$$

telles que, pour tout $t \in I$, $(\rho(t), \psi(t))$ soit un système de coordonnées polaires de $M(t)$:

$$\forall t \in I \quad \overrightarrow{OM(t)} = \rho(t)\,\overrightarrow{u}(\psi(t))$$

Démonstration : On passe en affixe complexe, en considérant l'application de classe C^k

$$z : I \to \mathbb{C}^* \quad t \mapsto z(t) = x(t) + i\, y(t)$$

si $(x(t), y(t))$ sont les coordonnécs de $M(t)$ dans $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$. Il suffit d'appliquer à la fonction $t \mapsto z(t)$ le théorème de relèvement **10-1**.57 pour obtenir une détermination $t \mapsto \psi(t)$ de classe C^k de son argument. Comme la fonction ρ définie par

$$\forall t \in I \quad \rho(t) = |z(t)| = \sqrt{x^2(t) + y^2(t)}$$

est composée de la fonction $t \mapsto x^2(t) + y^2(t)$ de classe C^k sur I, à valeurs dans $\mathbb{R}^{*+}$, et de $s \mapsto \sqrt{s}$ de classe C^∞ sur $\mathbb{R}^{*+}$, le couple de fonctions (ρ, ψ) répond évidemment à la question. ■

Sous les hypothèses précédentes, nous avons

$$\forall t \in I \quad \psi'(t) = \operatorname{Im} \frac{z'(t)}{z(t)} = \frac{x(t)\, y'(t) - y(t)\, x'(t)}{x^2(t) + y^2(t)}$$

DÉFINITION 20-1.25 *On dit qu'un arc paramétré Γ de classe C^k $I \ni t \mapsto M(t)$ admet une équation polaire de la forme $r = f(\theta)$ si et seulement s'il existe un intervalle $J \subset \mathbb{R}$ et une application $f : J \to \mathbb{R}$ de classe C^k tel que Γ soit C^k-équivalent à l'arc*

$$P : J \to \mathcal{E}_2 \quad \theta \mapsto P(\theta) \ \ \textit{vérifiant} \quad \overrightarrow{OP(\theta)} = f(\theta)\, \overrightarrow{u}(\theta)$$

En d'autres termes, on peut utiliser un paramétrage admissible $J \ni \theta \mapsto P(\theta)$ de l'arc géométrique associé à Γ tel que

$$\forall \theta \in J \quad (f(\theta), \theta) \text{ est un système de coordonnées polaires de } P(\theta)$$

La signification géométrique du paramètre θ va amener à utiliser des techniques particulières pour étudier de tels arcs. En particulier, on utilisera le repère mobile $(\overrightarrow{u}(\theta), \overrightarrow{v}(\theta))$ pour effectuer les calculs (tangente, concavité, asymptotes), de préférence au repère fixe $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$.

La proposition qui suit donne une condition nécessaire et suffisante pour qu'un arc de classe C^k dont le support ne contient pas l'origine O admette une équation polaire de la forme $r = f(\theta)$.

PROPOSITION 20-1.26 Soit $I \ni t \mapsto M(t)$ un arc de classe C^k dont le support ne contient pas O. Cet arc admet une équation polaire de la forme $r = f(\theta)$ si et seulement s'il est régulier et si aucune de ses tangentes ne passe par O.

Démonstration : On sait que la régularité et les tangentes sont des notions qui se conservent par changement de paramètre admissible. Si l'arc géométrique considéré admet une paramétrisation polaire

$$J \ni \theta \mapsto P(\theta) \quad \text{avec} \quad \overrightarrow{OP(\theta)} = f(\theta)\, \overrightarrow{u}(\theta)$$

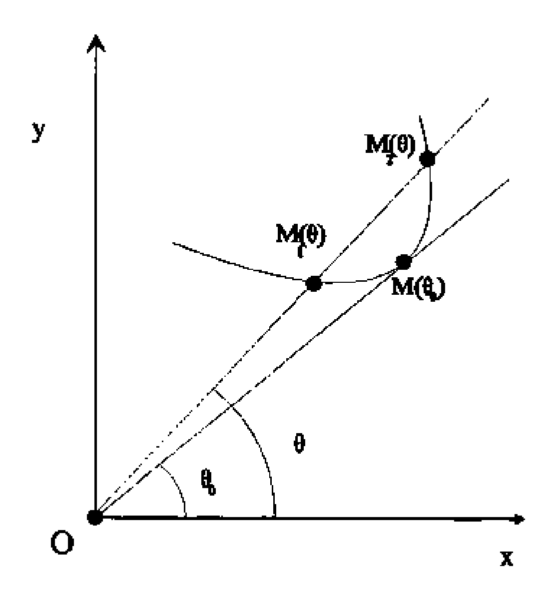


Figure **20.6** – *Arc n'ayant pas de représentation polaire*

avec f de classe C^k telle que $\forall\,\theta \in J \quad f(\theta) \neq 0$ (puisque $P(\theta) \neq O$), on a

$$\forall\,\theta \in J \quad \overrightarrow{P}'(\theta) = f'(\theta)\ \overrightarrow{u}(\theta) + f(\theta)\ \overrightarrow{v}(\theta) \neq \overrightarrow{0}$$

L'arc est donc régulier et $\overrightarrow{P}'(\theta)$ dirige la tangente en θ. De plus, comme $f(\theta) \neq 0$, les vecteurs $\overrightarrow{P}'(\theta)$ et $\overrightarrow{OP(\theta)}$ sont indépendants, ce qui montre que la tangente ne contient pas O.

Réciproquement, si l'arc $I \ni t \mapsto M(t)$ ne passe pas par O, on sait (proposition **20-1**.24) que l'on peut trouver deux fonctions ρ et ψ de classe C^k sur I telles que

$$\forall\,t \in I \quad \overrightarrow{OM(t)} = \rho(t)\ \overrightarrow{u}(\psi(t))$$

On a de plus, en notant $[\overrightarrow{x}, \overrightarrow{y}]$ le produit mixte des vecteurs $\overrightarrow{x}$ et $\overrightarrow{y}$

$$\forall\,t \in I \quad \psi'(t) = \frac{\left[\overrightarrow{OM}(t), \overrightarrow{M}'(t)\right]}{\left\|\overrightarrow{OM(t)}\right\|^2}$$

Si l'arc est régulier et si aucune de ses tangentes ne passe par O, nous avons

$$\forall\,t \in I \quad \psi'(t) \neq 0$$

et la fonction ψ réalise donc un C^k-difféomorphisme de I dans un intervalle J. Nous pouvons considérer cette application comme changement de paramètre admissible et, en considérant le difféomorphisme φ inverse

$$J \ni \theta \mapsto t = \varphi(\theta) = \psi^{-1}(\theta) \in I$$

nous obtenons une paramétrisation polaire de l'arc

$$\forall t \in I \quad M(t) = P(\theta)$$

$P(\theta) = M(\varphi(\theta))$ vérifiant

$$\forall \theta \in J \quad \overrightarrow{OP(\theta)} = \rho(\varphi(\theta))\,\overrightarrow{u}(\psi \circ \varphi(\theta)) = \rho(\varphi(\theta))\,\overrightarrow{u}(\theta)$$

L'arc considéré possède donc l'équation polaire $r = \rho \circ \varphi(\theta)$. ∎

20-1.5.3 Tangente, branches infinies, concavité

Nous considérons dans cette section un arc admettant une paramétrisation polaire

$$J \ni \theta \mapsto M(\theta) \quad \text{avec} \quad \overrightarrow{OM(\theta)} = r(\theta)\,\overrightarrow{u}(\theta)$$

où r est une fonction de classe $\mathcal{C}^k$ sur J. Rappelons que $r(\theta)$ doit être interprété comme une mesure algébrique et non comme une distance.

- **Tangentes** :

  On a déjà vu que

  $$\forall \theta \in J \quad \overrightarrow{M}'(\theta) = r'(\theta)\,\overrightarrow{u}(\theta) + r(\theta)\,\overrightarrow{v}(\theta)$$

  Toute valeur $\theta \in J$ avec $(r(\theta), r'(\theta)) \neq (0,0)$ est régulière :

  En un tel point, le vecteur $r'(\theta)\,\overrightarrow{u}(\theta) + r(\theta)\,\overrightarrow{v}(\theta)$ dirige la tangente $\mathcal{T}_\theta$ et le "rayon vecteur" $\overrightarrow{OM(\theta)}$ est colinéaire à $\overrightarrow{u}(\theta)$. Si on note $V(\theta)$ une détermination (définie modulo π) de l'angle orienté des droites $OM(\theta)$ et $\mathcal{T}_\theta$

  $$V(\theta) = (\widehat{OM(\theta), \mathcal{T}_\theta}) \pmod{\pi}$$

  nous avons

  $$\begin{cases} V(\theta) = \dfrac{\pi}{2} \pmod{\pi} & \text{si } r'(\theta) = 0 \\ \tan V(\theta) = \dfrac{r(\theta)}{r'(\theta)} & \text{sinon} \end{cases}$$

  On accède parfois mieux par le calcul à la cotangente de $V(\theta)$, qui est la dérivée logarithmique de la fonction r en θ :

  $$\operatorname{cotan} V(\theta) = \frac{r'(\theta)}{r(\theta)}$$

  L'arc ne peut présenter de point stationnaire que pour une valeur θ_0 avec $r(\theta_0) = 0$, c'est à dire pour lequel $M(\theta_0) = O$.

  En un tel point, la tangente est facile à placer : c'est la droite passant par O qui fait un angle θ_0 avec l'axe polaire. En effet, pour $\theta \to \theta_0$ la position limite de la sécante $OM(\theta)$ est bien la droite d'angle polaire θ_0. L'allure locale du support au voisinage de θ_0 dépend du signe de $r(\theta)$:

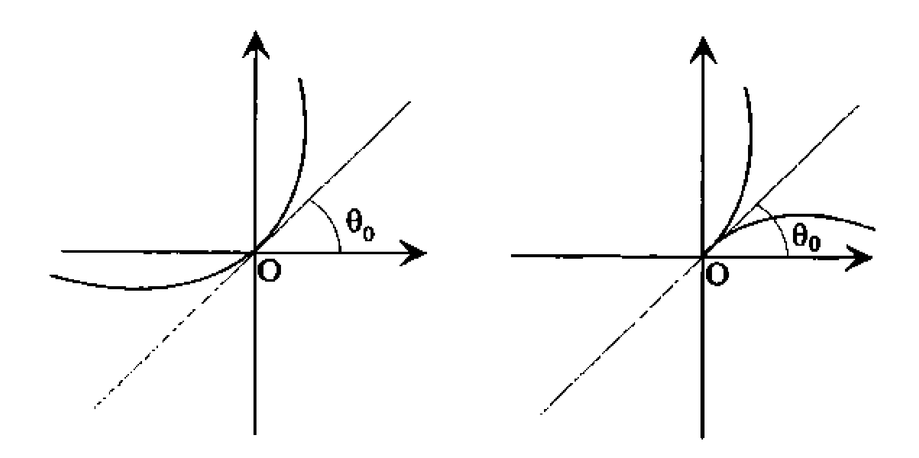


Figure **20.7** – *Allure à l'origine si $r(\theta)$ change ou non de signe en θ_0*

- Si $r(\theta)$ change de signe en θ_0 (c'est notamment le cas lorsque $r'(\theta_0) \neq 0$), on voit que le support de l'arc reste du même côté de sa tangente (point à concavité).
- Si $r(\theta)$ ne change pas de signe en θ_0 (qui est donc un point stationnaire), l'arc reste du même côté de la droite $O + \mathbb{R}\overrightarrow{v}(\theta_0)$, et on a l'aspect d'un point de rebroussement de première espèce.

- **Asymptotes : nous étudions ici le cas où $r(\theta) \to \pm\infty$ pour $\theta \to \theta_0$.**

  Pour θ tendant vers θ_0, la position limite de la droite $M(\theta)$ est évidemment la droite d'angle polaire θ_0. L'arc présente donc une direction asymptotique qui est celle de cette droite. Pour étudier l'existence éventuelle d'une asymptote, on efffectue un changement de repère[6]. On travaille dans $(O, \overrightarrow{u}(\theta_0), \overrightarrow{v}(\theta_0))$, le nouvel axe des abscisses étant parallèle à la direction asymptotique. L'arc présentera une asymptote si et seulement si l'ordonnée $Y(\theta)$ de $M(\theta)$ dans ce nouveau repère possède une limite pour θ tendant vers θ_0. Comme

$$Y(\theta) = \overrightarrow{OM}(\theta) \cdot \overrightarrow{v}(\theta_0) = r(\theta)\left(\overrightarrow{u}(\theta) \cdot \overrightarrow{v}(\theta_0)\right) = r(\theta)\sin(\theta - \theta_0)$$

[6]Sauf si la direction asymptotique est celle d'un des axes de coordonnées, c'est à dire lorsque

$$\theta_0 = k\frac{\pi}{2} \quad \text{avec} \quad k \in \mathbb{Z}$$

Dans ce cas, on reste dans le repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$ et on étudie

$$\lim_{\theta_0} x(\theta) = \lim_{\theta_0} r(\theta)\cos\theta$$

lorsque la direction asymptotique est verticale (k impair) ou

$$\lim_{\theta_0} y(\theta) = \lim_{\theta_0} r(\theta)\sin\theta$$

lorsque la direction asymptotique est horizontale.

l'arc présente une asymptote d'équation $Y = l$ (dans le repère local !) si et seulement si

$$\lim_{\theta \to \theta_0} r(\theta) \sin(\theta - \theta_0) = l$$

Si cette limite vaut $\pm\infty$, on aura seulement une branche parabolique dans la direction de la droite d'angle polaire θ_0.

Lorsqu'on a prouvé l'existence d'une asymptote, pour placer[7] le support de l'arc par rapport à cette asymptote, on doit étudier le signe de l'infiniment petit

$$\delta(\theta) = r(\theta) \sin(\theta - \theta_0) - l$$

Il est aussi indispensable de bien noter le signe de $r(\theta)$ pour θ proche de θ_0 : le point $M(\theta)$ s'éloigne à l'infini pour $\theta \to \theta_0$ mais, si $r(\theta) \to -\infty$, le vecteur $\overrightarrow{OM(\theta)}$ est de sens opposé à $\vec{u}(\theta)$.

- **Concavité :**

Si on suppose l'arc au moins de classe C^2, **les points non biréguliers** sont ceux qui vérifient

$$\left[\vec{M}'(\theta), \vec{M}''(\theta)\right] = 0$$

On évalue ce produit mixte dans la base directe $(\vec{u}(\theta), \vec{v}(\theta))$:

$$\vec{M}'(\theta) = r'(\theta)\ \vec{u}(\theta) + r(\theta)\ \vec{v}(\theta)$$
$$\text{et} \quad \vec{M}''(\theta) = (r''(\theta) - r(\theta))\ \vec{u}(\theta) + 2r'(\theta)\ \vec{v}(\theta)$$

On obtient donc

$$\left[\vec{M}'(\theta), \vec{M}''(\theta)\right] = \begin{vmatrix} r' & r'' - r \\ r & 2r' \end{vmatrix} = r^2(\theta) + 2r'^2(\theta) - r(\theta) r''(\theta)$$

- Lorsque la quantité $r^2(\theta) + 2r'^2(\theta) - r(\theta) r''(\theta)$ ne s'annule pas, le point considéré est un point birégulier, donc à concavité. Si P est un point quelconque de $\mathcal{E}_2$, on dit que l'arc tourne sa concavité en θ vers P si et seulement si P est dans la concavité de l'arc en θ. Pour que l'arc tourne sa concavité vers l'origine O, il faut et il suffit que les vecteurs $\overrightarrow{M(\theta)O}$ et $\vec{M}''(\theta)$ aient, avec un même vecteur normal à l'arc, des produits scalaires de même signe. En prenant comme vecteur normal

$$\vec{n} = -r(\theta)\ \vec{u}(\theta) + r'(\theta)\ \vec{v}(\theta)$$

la condition est $\left(\vec{M}''(\theta) . \vec{n}\right) \left(\overrightarrow{M(\theta)O} . \vec{n}\right) > 0$, soit

$$\left(r^2(\theta) + 2r'^2(\theta) - r(\theta) r''(\theta)\right) r^2(\theta) > 0$$

On a donc

$$r^2(\theta) + 2r'^2(\theta) - r(\theta) r''(\theta) > 0 \Leftrightarrow \text{ l'arc tourne sa concavité vers } O$$

[7] Attention : dans le repère local !

– On a de plus, puisque le produit mixte ne dépend pas de la base orthonormale directe utilisée pour écrire les déterminants,

$$r^2(\theta) + 2r'^2(\theta) - r(\theta)r''(\theta) = \left[\overrightarrow{M}'(\theta), \overrightarrow{M}''(\theta)\right] = \begin{vmatrix} x'(\theta) & x''(\theta) \\ y'(\theta) & y''(\theta) \end{vmatrix}$$

L'étude faite à la section **20-1.4** montre que, si

$$r^2(\theta) + 2r'^2(\theta) - r(\theta)r''(\theta)$$

s'annule en changeant de signe, alors θ_0 est un point d'inflexion de l'arc. Par contre, si cette quantité s'annule en θ_0 en gardant un signe constant, alors θ_0 n'est pas birégulier mais n'est pas un point d'inflexion.

– **Remarque :**
Un arc dont le support ne passe pas par l'origine peut être caractérisé par la fonction

$$q(\theta) = \frac{1}{r(\theta)}$$

et il est parfois préférable de travailler avec la fonction q plutôt qu'avec son inverse. On a alors

$$q'(\theta) = -\frac{r'(\theta)}{r^2(\theta)} \quad \text{et} \quad q''(\theta) = \frac{2r'^2(\theta) - r(\theta)\,r''(\theta)}{r^3(\theta)}$$

et donc

$$q(\theta) + q''(\theta) = \frac{r^2(\theta) + 2r'^2(\theta) - r(\theta)r''(\theta)}{r^3(\theta)}$$

Les points d'inflexion sont ceux pour lesquels $q(\theta) + q''(\theta)$ s'annule en changeant de signe. De même, l'arc tourne sa concavité vers l'origine O si la quantité $q(\theta)(q(\theta) + q''(\theta))$ est strictement positive (puique r^3 et q ont même signe).

20-1.5.4 Plan d'étude d'un arc défini par une équation polaire

On considère une fonction $\theta \mapsto r(\theta)$, et on étudie l'arc défini par

$$\theta \mapsto \overrightarrow{OM(\theta)} = r(\theta)\ \overrightarrow{u}(\theta)$$

- On commence par déterminer le domaine de définition $\mathcal{D}$ de la fonction r ainsi que sa régularité (classe C^k). Dans la pratique, $\mathcal{D}$ est souvent une réunion d'intervalles.

- On détermine ensuite **un sous-ensemble de description totale du support de l'arc** : on sait en effet que deux systèmes de coordonnées polaires (r_1, θ_1) et (r_2, θ_2) représentent le même point du plan (autre que O) si et seulement si

$$(r_1 = r_2 \text{ et } \theta_1 \equiv \theta_2 \mod 2\pi) \quad \text{ou} \quad (r_1 = -r_2 \text{ et } \theta_1 \equiv \theta_2 + \pi \mod 2\pi)$$

Il en résulte que, s'il existe un $k \in \mathbb{N}^*$ tel que

$$\forall \theta \in \mathbb{R} \quad \theta \in \mathcal{D} \Leftrightarrow \theta + 2k\pi \in \mathcal{D} \text{ et } r(\theta + 2k\pi) = r(\theta)$$

il suffira de faire parcourir à θ l'ensemble $\mathcal{D} \cap [0, 2k\pi]$ pour décrire la totalité du support de l'arc. On choisira évidemment k le plus petit possible dans $\mathbb{N}^*$. Si cette valeur de k est **impaire** et vérifie

$$\forall \theta \in \mathcal{D} \quad r(\theta + k\pi) = -r(\theta)$$

les points $M(\theta)$ et $M(\theta + k\pi)$ sont alors confondus, et on décrira entièrement le support avec θ dans $\mathcal{D} \cap [0, k\pi]$ (ou $\mathcal{D} \cap K$ où K est un segment quelconque de longueur $k\pi$).

- **On réduit ensuite éventuellement l'intervalle d'étude** en tenant compte des symétries. Par exemple, si $r(\theta) = r(-\theta)$, le point $M(-\theta)$ se déduit de $M(\theta)$ par symétrie par rapport à l'axe Ox. Plus généralement, s'il existe $\alpha \in \mathbb{R}$

$$\forall \theta \in \mathcal{D} \quad r(2\alpha - \theta) = r(\theta)$$

les points $M(\theta)$ et $M(2\alpha - \theta)$ sont symétriques par rapport à la droite passant par O et d'angle polaire α. Lorsque

$$\forall \theta \in \mathcal{D} \quad r(2\alpha - \theta) = -r(\theta)$$

ces points se déduisent l'un de l'autre par symétrie par rapport à la droite d'angle polaire $\alpha + \frac{\pi}{2}$. Il importe de garder une trace précise des transformations géométriques qui permettent de reconstituer, avec ordre et méthode, la totalité du support à partir de la portion obtenue sur l'intervalle d'étude.

- **Etude du signe de la fonction r sur l'intervalle d'étude.** Plus que le sens de variation, c'est le signe de cette fonction qui importe pour un tracé sommaire du support. C'est lui qui permet de placer, pour une valeur de θ, le point $M(\theta)$ d'un côté ou de l'autre de O sur la droite d'angle polaire θ. Les valeurs de θ qui annulent $r(\theta)$ permettent de placer les tangentes à l'origine. On place certains points remarquables, ainsi que les tangentes correspondantes. Rappelons que, lorsque $r'(\theta) = 0$, la tangente est orthogonale au rayon vecteur.

- **Etude des branches infinies, en particulier des asymptotes.**

- Eventuellement, si les calculs ne sont pas trop complexes, recherche des **points d'inflexion** et des **points doubles**. Pour trouver ces derniers, on cherche, dans un intervalle minimal de description totale du support, des valeurs θ et θ' vérifiant

$$(r(\theta) = r(\theta') \text{ et } \theta \equiv \theta' \bmod 2\pi) \quad \text{ou} \quad (r(\theta) = -r(\theta') \text{ et } \theta \equiv \theta' + \pi \bmod 2\pi)$$

- On fait ensuite un tracé sommaire du support résumant les études précédentes.

EXEMPLE 20-1.27 Etudier la courbe d'équation polaire

$$r(\theta) = \frac{\cos(2\theta)}{\cos\left(\theta + \frac{\pi}{6}\right)}$$

On obtient les résultats suivants :

- La fonction r est définie et C^∞ sur le domaine

$$\mathcal{D} = \bigcup_{k\in\mathbb{Z}} \left]\frac{\pi}{3} + k\pi, \frac{\pi}{3} + (k+1)\pi\right[$$

- La fonction r est 2π-périodique sur son domaine de définition, et vérifie de plus

$$\forall \theta \in \mathcal{D} \quad r(\theta + \pi) = -r(\theta)$$

On décrit entièrement le support de l'arc en faisant varier θ dans l'intersection de $\mathcal{D}$ et d'un intervalle d'amplitude π. On travaillera par exemple sur

$$\mathcal{D}' = \left[0, \frac{\pi}{3}\right[\cup \left]\frac{\pi}{3}, \pi\right]$$

Il n'y a pas de symétrie apparente lorsqu'on restreint l'étude à $\mathcal{D}'$.

- Signe de r : on a le tableau de signe suivant

θ	0		$\frac{\pi}{4}$		$\frac{\pi}{3}$		$\frac{3\pi}{4}$		π
$r(\theta)$	2	$+$	0	$-$	$-\infty$ ‖ $+\infty$	$+$	0	$-$	-2

L'origine est donc un point double, avec deux tangentes correspondant aux deux bissectrices des axes.

- Branche infinie pour $\theta \to \frac{\pi}{3}$: on effectue un développement limité de

$$r\left(\frac{\pi}{3} + h\right)\sin(h) = -\cos\left(\frac{2\pi}{3} + 2h\right) = \frac{1}{2} + \sqrt{3}h + o(h)$$

ce qui donne une asymptote d'équation $Y = \frac{1}{2}$ dans le repère

$$\left(O, \vec{u}\left(\frac{\pi}{3}\right), \vec{v}\left(\frac{\pi}{3}\right)\right)$$

Si un point M a une ordonnée Y dans ce repère, on a

$$Y = \overrightarrow{OM}.\vec{v}\left(\frac{\pi}{3}\right) = \overrightarrow{OM}.\left(-\frac{\sqrt{3}}{2}\vec{\imath} + \frac{1}{2}\vec{\jmath}\right)$$

et l'équation de l'asymptote, dans le repère $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$ est donc

$$\sqrt{3}x - y + 1 = 0$$

La courbe est en dessous de l'asymptote pour $\theta \to \frac{\pi}{3}$ par valeurs inférieures, au dessus pour $\theta \to \frac{\pi}{3}$ par valeurs supérieures. Voir le tracé figure **20.8**.

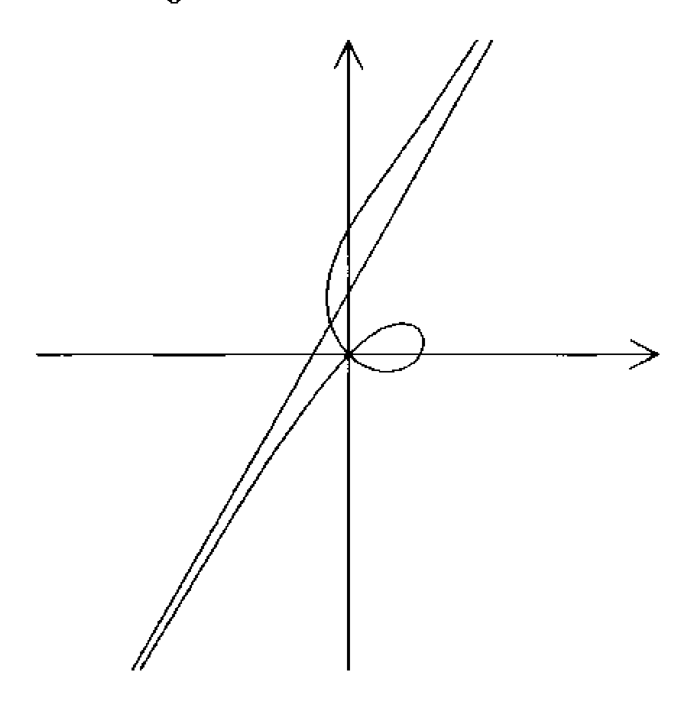


Figure **20.8** – *Courbe d'équation polaire* $r(\theta) = \dfrac{\cos(2\theta)}{\cos\left(\theta + \frac{\pi}{6}\right)}$

- Le tracé montre que le support de l'arc rencontre son asymptote. La valeur de θ correspondante est obtenue en résolvant $Y(\theta) = \frac{1}{2}$, soit

$$\cos(2\theta) = -\frac{1}{2} = \cos\left(\frac{2\pi}{3}\right)$$

Dans le domaine $\mathcal{D}'$, la solution distincte de $\frac{\pi}{3}$ est $\theta = \frac{2\pi}{3}$, ce qui donne le point d'intersection

$$Q\left(-\frac{\sqrt{3}}{6}, \frac{1}{2}\right)$$

EXERCICE 20-1.28 Etudier l'arc d'équation polaire

$$r(\theta) = \frac{\sin\frac{\theta}{2}}{\sin\frac{\theta}{2} + \cos\frac{\theta}{2}}$$

On déterminera en particulier l'asymptote, le point double et une valeur approchée des coordonnées du point d'inflexion.

EXERCICE 20-1.29 Etudier l'arc d'équation polaire

$$r(\theta) = a\cos\left(\frac{3\theta}{2}\right)$$

EXERCICE 20-1.30 Etudier l'arc d'équation polaire

$$r(\theta) = \frac{\sin\theta}{1+\cos\theta\cos 2\theta}$$

20-1.5.5 Quelques équations polaires remarquables

- **Droites :**

  Une équation d'une droite passant par le pôle O et d'angle polaire θ_0 est évidemment $\theta = \theta_0 + k\pi$, où $k \in \mathbb{Z}$ est quelconque. Si $\mathcal{D}$ est une droite du plan ne passant pas par O, elle admet dans le repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$ une équation de la forme $ux + vy + w = 0$, avec $(u,v) \neq (0,0)$ et $w \neq 0$. Si $M(x,y)$ est un point de $\mathcal{E}_2$ et si (r,θ) est un système de coordonnées polaires de M, on a

  $$M \in \mathcal{D} \Leftrightarrow r(u\cos\theta + v\sin\theta) = -w$$

  ce qui peut s'écrire

  $$r = \frac{1}{a\cos\theta + b\sin\theta}$$

  avec $(a,b) = \left(-\frac{u}{w}, -\frac{v}{w}\right) \neq (0,0)$. Réciproquement, toute équation de cette forme représente une droite ne passant pas par O, puisqu'elle peut s'écrire $ar\cos\theta + br\sin\theta = 1$, soit $ax + by - 1$ en coordonnées cartésiennes. Le vecteur

  $$\overrightarrow{n} = a\overrightarrow{i} + b\overrightarrow{j}$$

  est un vecteur normal à cette droite. Si (c,α) est un système de coordonnées polaires du point $A(a,b)$ avec, par exemple,

  $$c = \sqrt{a^2+b^2}, \quad \cos\alpha = \frac{a}{\sqrt{a^2+b^2}} \quad \text{et} \quad \sin\alpha = \frac{b}{\sqrt{a^2+b^2}}$$

  on a alors $\overrightarrow{n} = c\overrightarrow{u}(\alpha)$. Le vecteur $\overrightarrow{v}(\alpha)$ dirige donc la droite $\mathcal{D}$ (angle polaire $\alpha + \frac{\pi}{2}$), et l'équation peut s'écrire

  $$r = \frac{1}{c\cos(\theta - \alpha)}$$

  La projection orthogonale de O sur $\mathcal{D}$ est donc le point Ω tel que[8]

  $$\overrightarrow{O\Omega} = \frac{1}{c}\overrightarrow{u}(\alpha)$$

[8]Ce point est donc image du point $A(a,b)$ par inversion de pôle O et de puissance 1.

EXERCICE 20-1.31 Soit un arc de classe C^k admettant une équation polaire de la forme $r = f(\theta)$. Montrer que, si $f(\theta_0) \neq 0$, une équation polaire de la tangente à cet arc en θ_0 peut s'écrire

$$\frac{1}{r} = \frac{1}{f(\theta_0)} \cos(\theta - \theta_0) + \left(\frac{1}{f}\right)'(\theta_0) \sin(\theta - \theta_0)$$

Montrer que, si $f'(\theta_0) \neq 0$, une équation de la normale est

$$\frac{1}{r} = \frac{1}{f(\theta_0)} \cos(\theta - \theta_0) + \frac{1}{f'(\theta_0)} \sin(\theta - \theta_0)$$

- **Cercles passant par l'origine :**

Si $\mathcal{C}$ est un cercle de centre $C(a_1, b_1)$ passant par O, il possède une équation cartésienne de la forme

$$x^2 + y^2 - 2a_1 x - 2b_1 y = 0$$

(l'équation ne contient pas de terme constant, puisque le cercle passe par O). Comme précédemment, un point M de coordonnées polaires (r, θ) appartient au cercle si et seulement si

$$r^2 = (2a_1 \cos\theta + 2b_1 \sin\theta)\, r$$

En posant $(a, b) = (2a_1, 2b_1)$, le point $A(a, b)$ vérifie $\overrightarrow{OA} = 2\overrightarrow{OC}$ et $[O, A]$ est donc un diamètre du cercle. La condition qui précède peut s'écrire

$$r = 0 \quad \text{ou} \quad r = a\cos\theta + b\sin\theta$$

Comme la quantité $a\cos\theta + b\sin\theta$ s'annule pour certaines valeurs de θ, la totalité du cercle est donc décrite par l'équation polaire

$$r = a\cos\theta + b\sin\theta = \sqrt{a^2 + b^2}\cos(\theta - \alpha) = 2R\cos(\theta - \alpha)$$

où, comme précédemment, $\left(c = \sqrt{a^2 + b^2}, \alpha\right)$ est un système de coordonnées polaires du point A : l'angle polaire de la droite OA est α et le rayon du cercle est $R = \frac{1}{2}\sqrt{a^2 + b^2}$. Il est évident que, réciproquement, toute équation de cette forme est équation d'un cercle passant par l'origine. On remarquera que, pour le cercle comme pour la droite, l'expression de $r(\theta)$ obtenue vérifie $r(\theta + \pi) = -r(\theta)$, et qu'il suffit donc de faire parcourir à θ un intervalle de longueur π pour décrire la totalité de la droite ou du cercle (figure page 980).

EXERCICE 20-1.32 Expliquer comment le support de l'arc d'équation polaire

$$r(\theta) = a(1 + \cos\theta)$$

peut se construire géométriquement à partir du cercle d'équation $r = a\cos\theta$. Tracer le support de cet arc. On obtient une courbe appelée cardioïde, présentant un point de rebroussement en O. (Attention : pour décrire la totalité du support de l'arc, il faut faire varier (par exemple) θ dans $[0, 2\pi]$, et le cercle est alors parcouru deux fois, avec le rayon vecteur qui change de signe pour chaque passage par O).

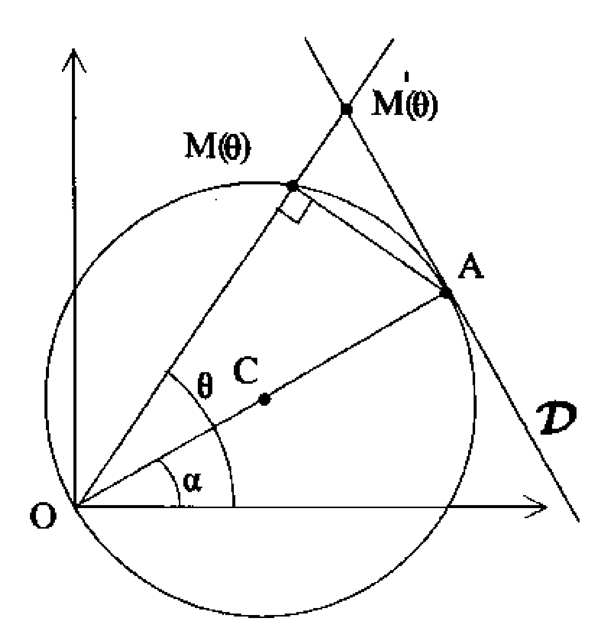


Figure **20.9** – *Equation polaire d'un cercle :* $\overline{OM(\theta)} = OA\cos(\theta-\alpha)$

EXERCICE 20-1.33 A l'exercice **18-2.7**, nous avons défini l'inversion de pôle O et de puissance k. C'est l'application de $\mathcal{E}_n - \{O\}$ dans lui même, définie par $M \mapsto M'$ tel que O, M et M' soient alignés avec

$$\overrightarrow{OM}.\overrightarrow{OM'} = k$$

Dans le plan complexe, montrer que cette application peut se traduire par la transformation des affixes

$$0 \neq z \mapsto Z = \frac{k}{\bar{z}}$$

Il est évident que l'inversion de puissance k se déduit de l'inversion de puissance 1 par composition avec l'homothétie de centre O et de rapport k.

En utilisant les équations polaires, montrer que l'image par inversion de pôle O d'une droite ne passant pas par O est un cercle passant par O (privé de O). Cette propriété est géométriquement évidente : sur la figure **20.9**, on considère le cercle $\mathcal{C}$ de diamètre $[O, A]$ et l'inversion de pôle O et de puissance $a^2 = OA^2$. Le point A est alors invariant, et si $\mathcal{D}$ est la perpendiculaire à OA passant par A on note M' l'intersection de la droite OM avec $\mathcal{D}$, M étant un point quelconque de $\mathcal{C} - \{O\}$. On a clairement

$$OM = OA\cos(\theta - \alpha) \quad \text{et} \quad OM' = \frac{OA}{\cos(\theta - \alpha)}$$

ce qui donne

$$OM.OM' = OA^2$$

et M' est bien image de M par l'inversion considérée. L'image du cercle par l'inversion de puissance k est alors une droite, image de $\mathcal{D}$ par l'homothétie de centre O et de rapport $\frac{k}{a^2}$. Cette droite est encore perpendiculaire à la droite OA. Quelle est l'image d'un cercle passant par O ?

REMARQUE 20-1.34 En effectuant des rotations d'axe OA, on obtient un résultat analogue en dimension 3 : l'image d'une sphère de diamètre $[O, A]$ par une inversion de pôle O est un plan perpendiculaire à la droite OA. Lorsque la puissance de l'inversion est OA^2, c'est le plan tangent à la sphère en A. L'application $M \mapsto M'$ de la sphère sur le plan est alors connue sous le nom de **projection stéréographique** de pôle O (figure **20.10**). Elle est utilisée en cartographie pour représenter sur un plan les régions de la sphère proches de A. L'intérêt de ce genre de projection est qu'elle **conserve les angles** : deux arcs tracés sur la sphère se coupant en faisant un angle (non orienté) α se projettent en deux arcs plans ayant la même propriété (voir toujours l'exercice **18-2**.7).

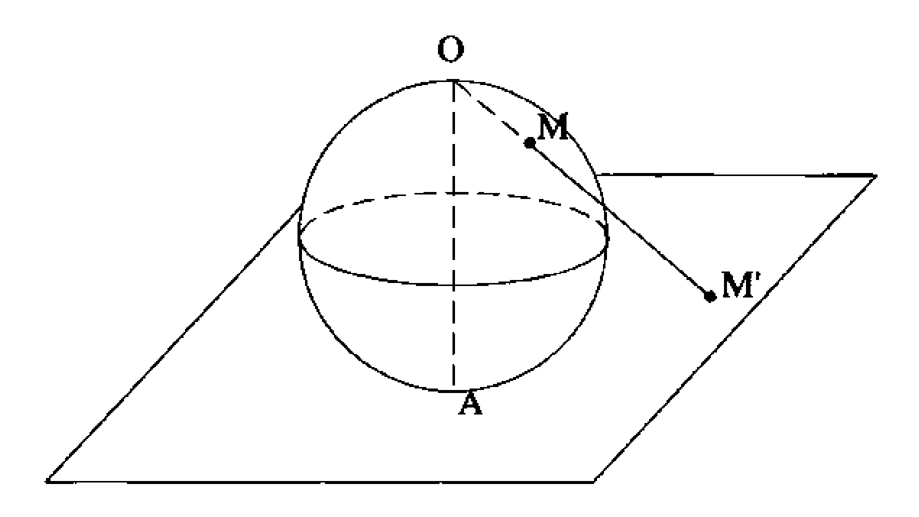


Figure **20.10** – *Projection stéréographique*

EXERCICE 20-1.35 Puissance d'un point par rapport à un cercle.

Soit C un cercle du plan euclidien, de centre Ω et de rayon R, et soit O un point n'appartenant pas à C. On considère une droite $\mathcal{D}$ passant par O et rencontrant le cercle en deux points M et M' (éventuellement confondus si O est extérieur au cercle). Montrer que le produit scalaire $\overrightarrow{OM}.\overrightarrow{OM'}$ vérifie

$$\overrightarrow{OM}.\overrightarrow{OM'} = O\Omega^2 - R^2$$

et ne dépend donc pas de la droite $\mathcal{D}$ choisie (on pourra faire intervenir la projection orthogonale H de Ω sur $\mathcal{D}$). On définit donc la puissance du point O par rapport au cercle C (notée $P_C(O)$) par

$$P_C(O) = \overrightarrow{OM}.\overrightarrow{OM'} = O\Omega^2 - R^2$$

Quelle est l'image du cercle C par l'inversion de pôle O et de puissance $P_C(O)$? En déduire que l'image par inversion d'un cercle ne passant pas par le pôle est également un cercle. Retrouver ce résultat par un calcul en affixe complexe.

EXERCICE 20-1.36 Donner une équation polaire de l'hyperbole équilatère d'équation

$$x^2 - y^2 = a^2$$

dans le repère orthonormé $\left(O,\vec{\imath},\vec{\jmath}\right)$. Déterminer l'image de cette hyperbole par l'inversion de pôle O et de puissance k (on trouve une courbe appelée lemniscate de Bernoulli).

- **Coniques de foyer O :**

  Voir la section **20-1.7.6**.

20-1.6 Définition d'une courbe plane par équation ou par paramétrage

Le plan $\mathcal{E}_2$ est rapporté à un repère $\mathcal{R}=\left(O,\vec{\imath},\vec{\jmath}\right)$. Pour définir une "courbe" du plan, nous disposons de deux méthodes, qui vont s'avérer localement équivalentes :

- **Définition par équation :**

  Soient U un ouvert de $\mathbb{R}^2$, et $\mathcal{U}=\{M(x,y)\in\mathcal{E}_2 \mid (x,y)\in U\}$ l'ouvert du plan correspondant. Si $F:U\to\mathbb{R}$ une fonction numérique de classe C^k,

$$\Gamma_F=\{M(x,y)\in\mathcal{U}\mid F(x,y)=0\}$$

  est l'ensemble des points du plan vérifiant l'équation cartésienne

$$F(x,y)=0$$

  Cet ensemble peut ne pas être ce qu'on appelle habituellement une "courbe" : si f est la fonction définie sur $\mathbb{R}$ par

$$\begin{cases} f(t)=0 & \text{si} \quad t\notin\,]1,4[\\ f(t)=e^{-\left(\frac{1}{(4-t)^2}+\frac{1}{(t-1)^2}\right)} & \text{sinon}\end{cases}$$

  $f\in C^\infty(\mathbb{R},\mathbb{R})$ (voir exercice **8-5**.20). La fonction $F:\mathbb{R}^2\to\mathbb{R}$ définie par

$$F(x,y)=f\left(x^2+y^2\right)$$

  est donc de classe C^∞ sur $\mathbb{R}^2$. L'ensemble Γ_F est clairement

$$\Gamma_F=\left\{M\in\mathcal{E}_2\mid \left\|\overrightarrow{OM}\right\|\leqslant 1 \quad \text{ou} \quad \left\|\overrightarrow{OM}\right\|\geqslant 2\right\}$$

  c'est à dire le complémentaire dans $\mathcal{E}_2$ d'une couronne circulaire ouverte.

  Par contre, l'étude de la section **18-3.4.3** nous donne le théorème (en supposant toujours F de classe C^k) :

THÉORÈME 20-1.37 Si $M_0(x_0,y_0)\in\Gamma_F$ est un point régulier (c'est à dire si une au moins des dérivées partielles de F est non nulle en M_0), il existe un voisinage ouvert $\mathcal{V}$ de M_0 tel que $\Gamma_F\cap\mathcal{V}$ soit le support

d'un arc paramétré de classe C^k. Lorsque $\dfrac{\partial F}{\partial y}(M_0) \neq 0$, l'abscisse x est un paramètre admissible pour cet arc. L'arc est donc régulier en x_0 et l'équation de la tangente en M_0 est

$$\frac{\partial F}{\partial x}(M_0)\,(x - x_0) + \frac{\partial F}{\partial y}(M_0)\,(y - y_0) = 0$$

Si l'espace est euclidien, le vecteur $\overrightarrow{\mathrm{grad}}F(M_0)$ dirige la normale à l'arc en M_0.

- **Définition par paramétrage :**

  Réciproquement, le support d'un arc de classe C^k défini par $I \ni t \mapsto M(t)$ admet, **au voisinage d'un point t_0 régulier**, une équation cartésienne : si $(x(t), y(t))$ sont les coordonnées de $M(t)$, et si $\overrightarrow{M}'(t_0) \neq \overrightarrow{0}$, on a par exemple $x'(t_0) \neq 0$. Par continuité, on peut trouver un réel $\alpha > 0$ tel que

$$\forall t \in]t_0 - \alpha, t_0 + \alpha[\quad x'(t) \neq 0$$

  Cette dérivée garde donc un signe constant sur $]t_0 - \alpha, t_0 + \alpha[$, et l'application $t \mapsto x(t)$ induit un C^k-difféomorphisme (donc un changement de variable admissible) entre $]t_0 - \alpha, t_0 + \alpha[$ et un intervalle ouvert J_0, voisinage de $x_0 = x(t_0)$. Si on note

$$J_0 \ni x \mapsto \varphi(x) = t \in]t_0 - \alpha, t_0 + \alpha[$$

  le difféomorphisme réciproque, pour tout $t \in]t_0 - \alpha, t_0 + \alpha[$ les coordonnées $(X, Y) = (x(t), y(t))$ de $M(t)$ vérifient l'équation cartésienne

$$F(X, Y) = Y - y(\varphi(X)) = 0$$

  avec F de classe C^k définie sur $J_0 \times \mathbb{R}$. Réciproquement, tout point $P(X, Y)$ dont les coordonnées sont dans $J_0 \times \mathbb{R}$ et vérifient cette équation est évidemment $P = M(t)$, avec $t = \varphi(X)$. On a de plus

$$\frac{\partial F}{\partial Y}(P) = 1$$

Il y a donc, lorsqu'on travaille avec des courbes et des arcs réguliers, équivalence locale entre définition par paramétrage et par équation.

EXEMPLE 20-1.38 Trouver un paramétrage rationnel de la courbe Γ d'équation

$$F(x, y) = x^3 - 8y^3 - 2xy - 2x + 4y - 3 = 0$$

Γ est l'ensemble des points du plan dont les coordonnées annulent un polynôme de deux variables, de degré 3. On dit que Γ est une **courbe algébrique** de degré 3. Si on coupe Γ par une droite variable $\mathcal{D}$ d'équation $y = \alpha x + \beta$, l'équation aux abscisses des points communs à Γ et $\mathcal{D}$ est en général une équation du troisième degré en x, qui peut donc posséder 3 racines réelles. S'il existe un point $M_0(x_0, y_0) \in \Gamma$ par lequel passent deux branches localement distinctes de Γ (on

dit alors que M_0 est un **point double** pour Γ), en faisant tourner autour de M_0 une droite $\mathcal{D}_t$ d'équation $y - y_0 = t\,(x - x_0)$, l'équation du troisième degré possèdera x_0 pour racine double, et on pourra donc expliciter la troisième racine $x = x\,(t)$, ce qui donnera les coordonnées du point d'intersection $M\,(t) = \Gamma \cap \mathcal{D}_t$

$$M\,(t) : \;(x\,(t)\,, y\,(t) = y_0 + t\,(x\,(t) - x_0))$$

On obtiendra ainsi un arc dont le support sera la totalité de Γ (si on omet les points M pour lesquels la droite M_0M est verticale, qu'on obtiendra probablement comme points limites en faisant tendre t vers $\pm\infty$).

Comment trouver les coordonnées du point double, s'il existe? En remarquant que, si M_0 existe, le théorème des fonctions implicites ne peut pas s'appliquer en M_0, et donc forcément

$$\frac{\partial F}{\partial x}\,(M_0) = \frac{\partial F}{\partial y}\,(M_0) = 0 \quad \text{avec évidemment } F\,(M_0) = 0$$

Ces conditions sur les dérivées donnent ici

$$\begin{cases} 3x_0^2 - 2y_0 - 2 = 0 \\ -x_0 - 12y_0^2 + 2 = 0 \end{cases}$$

En additionnant ces deux équations, on obtient

$$(x_0 + 2y_0)\,(3\,(x_0 - 2y_0) - 1) = 0$$

On essaie $x_0 = -2y_0$, et la première équation donne alors $3x_0^2 + x_0 - 2 = 0$. On vérifie ensuite que le point $M_0\left(-1, \frac{1}{2}\right)$ convient car il vérifie $F\,(M_0) = 0$. On obtiendra le paramétrage

$$x\,(t) = \frac{8t^3 + 12t^2 + 2t + 2}{1 - 8t^3} \quad \text{et} \quad y\,(t) = \frac{1}{2}\frac{16t^3 + 4t^2 + 6t + 1}{1 - 8t^3}$$

ce qui permet un tracé précis de la courbe à l'aide de l'ordinateur.

A l'inverse, trouver une équation cartésienne pour décrire le support d'un arc est parfois fort utile, pour éviter des études complexes. Le problème se ramène alors à "éliminer le paramètre t entre $x\,(t)$ et $y\,(t)$", c'est à dire trouver une condition nécessaire (et suffisante si possible) portant sur les coordonnées X et Y d'un point P du plan pour qu'il existe $t \in I$ avec $M\,(t) = P$ soit

$$\begin{cases} x\,(t) = X \\ y\,(t) = Y \end{cases}$$

EXERCICE 20-1.39 Tracer le support de l'arc défini dans un repère orthonormé du plan par

$$x\,(t) = \frac{t}{1 + t^4} \quad \text{et} \quad y\,(t) = \frac{t^3}{1 + t^4}$$

On montrera que le support de l'arc est une courbe algébrique de degré 4, qui admet une représentation polaire très simple (voir exercice **20-1**.36).

EXERCICE 20-1.40 On continue l'étude de l'arc précédent. Donner une condition nécessaire et suffisante sur t_1, t_2, t_3 et t_4 (supposés distincts) pour que les points $(M_i(t_i))_{1\leqslant i\leqslant 4}$ soient alignés. Lorsque la tangente en un point de paramètre t recoupe l'arc en deux points distincts de paramètres t' et t'', déterminer le centre du cercle circonscrit au triangle $OM(t')M(t'')$.

Indication : pour la première question, écrire *a priori* l'équation d'une droite Δ : $ux+vy+w=0$ et montrer que les points d'intersection de Δ et du support de l'arc correspondent à des valeurs du paramètre t qui sont racines d'un polynôme de degré 4. Si les quatre racines sont $(t_i)_{1\leqslant i\leqslant 4}$, utiliser les relations entre coefficients et racines d'un polynôme scindé. On obtiendra la condition

$$t_1t_2t_3t_4=1 \quad \text{et} \quad t_1t_2+t_1t_3+t_1t_4+t_2t_3+t_2t_4+t_3t_4=0$$

Pour la deuxième question, en utilisant la notion de contact et ce qui précède, déterminer $(t'+t'')$ et $t't''$. Ecrire ensuite *a priori* l'équation d'un cercle passant par O, et voir à quelle condition ce cercle recoupe l'arc en $M(t')$ et $M(t'')$. Le centre du cercle cherché est $M(-t)$.

20-1.7 Courbes du second degré et coniques

20-1.7.1 Courbe du second degré

Nous travaillons ici dans un plan affine $\mathcal{E}_2$ basé sur l'espace vectoriel E_2.

DÉFINITION 20-1.41 *Le plan $\mathcal{E}_2$ affine étant ramené à un repère quelconque $\left(O,\vec{\imath},\vec{\jmath}\right)$, on appelle courbe du second degré tout sous-ensemble du plan caractérisé par une équation polynomiale de degré 2*

$$\Gamma=\{M(x,y)\in\mathcal{E}_2 \mid Q(x,y)=0\}$$

où Q est un polynôme des deux variables x et y,

$$Q(x,y)=A\,x^2+2B\,xy+C\,y^2+2D\,x+2E\,y+F$$

avec $(A,B,C,D,E,F)\in\mathbb{R}^6$ et $(A,B,C)\neq 0$

Si Φ est la forme quadratique sur E_2 et l est la forme linéaire dont les matrices dans la base $\left(\vec{\imath},\vec{\jmath}\right)$ sont respectivement

$$M=\begin{pmatrix} A & B \\ B & C \end{pmatrix} \quad \text{et} \quad L=(D,F)$$

on a donc

$$M\in\Gamma \Leftrightarrow \Phi\left(\overrightarrow{OM}\right)+2l\left(\overrightarrow{OM}\right)+F=0$$

ce qui montre que, dans tout repère $\left(O,\vec{\imath}_1,\vec{\jmath}_1\right)$ d'origine O, la courbe Γ possède une équation du second degré, les matrices de Φ et l se transformant selon la règle

habituelle : si P est la matrice de passage de $\left(\vec{\imath}, \vec{\jmath}\right)$ à $\left(\vec{\imath}_1, \vec{\jmath}_1\right)$, l'équation de Γ dans $\left(O, \vec{\imath}_1, \vec{\jmath}_1\right)$ sera de la forme

$$A_1 x^2 + 2B_1 xy + C_1 y^2 + 2D_1 x + 2E_1 y + F = 0$$

avec

$$M_1 = \begin{pmatrix} A_1 & B_1 \\ B_1 & C_1 \end{pmatrix} = {}^tPMP \quad \text{et} \quad L_1 = (D_1, E_1) = LP$$

De même, si on change l'origine du repère, nous aurons pour $\Omega \in \mathcal{E}_2$ quelconque

$$M \in \Gamma \Leftrightarrow \Phi\left(\overrightarrow{O\Omega} + \overrightarrow{\Omega M}\right) + 2l\left(\overrightarrow{O\Omega} + \overrightarrow{\Omega M}\right) + F = 0$$

condition qui peut encore s'écrire

$$\Phi\left(\overrightarrow{\Omega M}\right) + 2l_1\left(\overrightarrow{\Omega M}\right) + F_1 = 0$$

où l_1 est la forme linéaire

$$l_1 = l + \varphi\left(\overrightarrow{O\Omega}, \bullet\right)$$

et où φ désigne la forme polaire de Φ, avec $F_1 = F + \Phi\left(\overrightarrow{O\Omega}\right) + 2l\left(\overrightarrow{O\Omega}\right)$. **Dans le repère $\left(\Omega, \vec{\imath}_1, \vec{\jmath}_1\right)$, Γ possède encore une équation du second degré, qui fait intervenir la même forme quadratique Φ.**

20-1.7.2 Réduction de l'équation en repère orthonormé

On suppose à présent le plan $\mathcal{E}_2$ euclidien, et le repère $\left(O, \vec{\imath}, \vec{\jmath}\right)$ orthonormé. On cherche un autre repère orthonormé $\left(\Omega, \vec{\imath}_1, \vec{\jmath}_1\right)$ où l'équation sera plus simple (pas de terme en xy ni en général de terme du premier degré). Ceci permettra de décrire géométriquement Γ. La discussion se fait selon la signature de la forme quadratique Φ :

Comme $\mathcal{B} = \left(\vec{\imath}, \vec{\jmath}\right)$ est une base orthonormale, la matrice de Φ dans cette base est aussi matrice de l'endomorphisme symétrique u associé à Φ :

$$M(u, \mathcal{B}) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

et nous savons (*cf.* théorème **14-3**.28) qu'il existe une base orthonormale

$$\mathcal{B}_1 = \left(\vec{\imath}_1, \vec{\jmath}_1\right)$$

diagonalisant u. Si (λ, μ) sont les valeurs propres associées, la matrice de Φ dans cette base est

$$M(\Phi, \mathcal{B}_1) = M(u, \mathcal{B}_1) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

La signature de Φ est donc donnée par les signes des valeurs propres de u. Le déterminant de u vaut

$$\det u = AC - B^2 = \lambda\mu$$

1. $\boxed{AC - B^2 > 0}$: **la signature de Φ est $(2,0)$ ou $(0,2)$. On dit que Γ est une courbe du second degré du genre ellipse.**

Les valeurs propres λ et μ sont de même signe. En changeant éventuellement l'équation en son opposée, nous supposerons

$$0 < \lambda \leqslant \mu$$

Dans le repère $\left(O, \overrightarrow{I}_1, \overrightarrow{J}_1\right)$, où les coordonnées du point générique M sont notées (X_1, Y_1), la caractérisation des points de Γ devient

$$\begin{aligned} M(X_1, Y_1) \in \Gamma \Leftrightarrow \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + F \\ = \lambda X_1^2 + \mu Y_1^2 + 2\alpha X_1 + 2\beta Y_1 + F = 0 \end{aligned}$$

On peut alors faire disparaître les termes du premier degré, en les faisant entrer dans des carrés :

$$\begin{aligned} \lambda X_1^2 + \mu Y_1^2 + 2\alpha X_1 + 2\beta Y_1 + F \\ = \lambda\left(X_1 + \frac{\alpha}{\lambda}\right)^2 + \mu\left(Y_1 + \frac{\beta}{\mu}\right)^2 + F - \frac{\alpha^2}{\lambda} - \frac{\beta^2}{\mu} \end{aligned}$$

En choisissant comme nouvelle origine le point Ω dont les coordonnées dans $\left(O, \overrightarrow{I}_1, \overrightarrow{J}_1\right)$ sont $\left(-\frac{\alpha}{\lambda}, -\frac{\beta}{\mu}\right)$, les coordonnées (X, Y) du point générique M dans $\left(\Omega, \overrightarrow{I}_1, \overrightarrow{J}_1\right)$ sont

$$\begin{cases} X = X_1 + \dfrac{\alpha}{\lambda} \\ Y = Y_1 + \dfrac{\beta}{\mu} \end{cases}$$

et donc

$$M(X, Y) \in \Gamma \Leftrightarrow \lambda X^2 + \mu Y^2 + f' = 0$$

avec $f' = F - \frac{\alpha^2}{\lambda} - \frac{\beta^2}{\mu}$. La nature exacte de Γ dépend alors du signe de f' :

- si $f' > 0$, Γ est réduit à **l'ensemble vide.**
- Si $f' = 0$, $M(X, Y) \in \Gamma \Leftrightarrow X = Y = 0$, et donc Γ **se réduit à un point** : $\Gamma = \{\Omega\}$.
- Si $f' < 0$, en divisant l'équation par $-f'$, et en posant

$$a^2 = -\frac{f'}{\lambda} \quad \text{et} \quad b^2 = -\frac{f'}{\mu}$$

on obtient l'équation réduite de Γ

$$M(X, Y) \in \Gamma \Leftrightarrow \frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$$

La courbe Γ est donc **une ellipse de centre** Ω et de grand axe égal à $\Omega+\text{vect}\left(\overrightarrow{I}\right)$ (où $\overrightarrow{I}$ est vecteur propre de u associé à la plus petite valeur propre en valeur absolue).

Dans le cas particulier $\lambda=\mu$, on trouve donc un cercle (éventuellement vide ou réduit à un point), mais cela apparaissait en fait déjà dans l'équation de départ (en repère orthonormé) : si $\lambda=\mu$, on a en fait

$$u=\lambda\, id_{E_2}$$

et donc, dans toute base orthonormale de E_2 l'expression de Φ est $\lambda\left(x^2+y^2\right)$ (pas de terme en xy), puisque

$$\forall\, \overrightarrow{x}\in E_2\quad \Phi\left(\overrightarrow{x}\right)=\lambda\left\|\overrightarrow{x}\right\|^2$$

2. $\boxed{AC-B^2<0}$: **la signature de Φ est** $(1,1)$. On dit que Γ **est une courbe du second degré du genre hyperbole.**

Les valeurs propres λ et μ sont de signes opposés. En permutant éventuellement les vecteurs propres de u, nous supposerons

$$\mu<0<\lambda$$

Dans le repère $\left(O,\overrightarrow{I}_1,\overrightarrow{J}_1\right)$, où les coordonnées du point générique M sont notées (X_1,Y_1), la caractérisation des points de Γ devient

$$\begin{aligned}M\left(X_1,Y_1\right)\in\Gamma\Leftrightarrow \Phi\left(\overrightarrow{OM}\right)+2l\left(\overrightarrow{OM}\right)+F&\\ =\lambda X_1^2+\mu Y_1^2+2\alpha X_1+2\beta Y_1+F&=0\end{aligned}$$

Comme dans le cas de l'ellipse, on écrit :

$$\begin{aligned}\lambda X_1^2+\mu Y_1^2+2\alpha X_1+2\beta Y_1+F&\\ =\lambda\left(X_1+\frac{\alpha}{\lambda}\right)^2+\mu\left(Y_1+\frac{\beta}{\mu}\right)^2+F-\frac{\alpha^2}{\lambda}-\frac{\beta^2}{\mu}&\end{aligned}$$

Le point Ω ayant toujours les coordonnées $\left(-\dfrac{\alpha}{\lambda},-\dfrac{\beta}{\mu}\right)$ dans $\left(O,\overrightarrow{I}_1,\overrightarrow{J}_1\right)$, l'équation de Γ dans $\left(\Omega,\overrightarrow{I}_1,\overrightarrow{J}_1\right)$ est

$$M\left(X,Y\right)\in\Gamma\Leftrightarrow\lambda X^2+\mu Y^2+f'=0$$

avec $f'=F-\dfrac{\alpha^2}{\lambda}-\dfrac{\beta^2}{\mu}$. La nature exacte de Γ varie selon que $f'=0$ ou $f'\neq 0$:

- Si $f'=0$, $M\left(X,Y\right)\in\Gamma\Leftrightarrow\lambda X^2=-\mu Y^2$, et donc Γ **est réunion de deux droites concourantes en** Ω, d'équations

$$Y=\pm\sqrt{-\frac{\lambda}{\mu}}X$$

(les directions de ces droites sont les directions des deux droites vectorielles qui annulent la forme quadratique Φ)

- Si $f' \neq 0$, en divisant l'équation par $|f'|$, et en posant

$$a^2 = \frac{|f'|}{\lambda} \quad \text{et} \quad b^2 = -\frac{|f'|}{\mu}$$

on obtient l'équation réduite de Γ

$$M(X,Y) \in \Gamma \Leftrightarrow \frac{X^2}{a^2} - \frac{Y^2}{b^2} = \pm 1$$

La courbe Γ est donc **une hyperbole de centre** Ω et d'axes

$$\mathcal{D}_1 = \Omega + \text{vect}\left(\vec{I}\right) \text{ et } \mathcal{D}_2 = \Omega + \text{vect}\left(\vec{J}\right)$$

(l'axe focal est $\mathcal{D}_1$ si le second membre de l'équation réduite est $+1$, $\mathcal{D}_2$ sinon). Les **asymptotes ont pour équations**, dans $\left(\Omega, \vec{I}_1, \vec{J}_1\right)$

$$Y = \pm\frac{b}{a}X$$

et sont dirigées par les vecteurs qui annulent la forme quadratique Φ. L'hyperbole est **équilatère** si et seulement si $a = b$, soit $\lambda = -\mu$, ce qui peut aussi s'écrire

$$A + C = 0$$

puisque $\lambda + \mu = \text{trace } u = A + C$.

3. $\boxed{AC - B^2 = 0}$: **Φ est dégénérée (signature $(1,0)$ ou $(0,1)$). Γ est du genre parabole.**

L'endomorphisme u possède alors les valeurs propres $\lambda \neq 0$ et $\mu = 0$ (ce qui revient aussi à dire que Φ est, au signe près, la carré d'une forme linéaire).

Dans le repère $\left(O, \vec{I}_1, \vec{J}_1\right)$, où les coordonnées du point générique M sont notées (X_1, Y_1), la caractérisation des points de Γ s'écrit

$$M(X_1, Y_1) \in \Gamma \Leftrightarrow \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + F$$
$$= \lambda X_1^2 + 2\alpha X_1 + 2\beta Y_1 + F = 0$$

Comme dans le cas de l'ellipse, on écrit :

$$\lambda X_1^2 + 2\alpha X_1 + 2\beta Y_1 + F = \lambda\left(X_1 + \frac{\alpha}{\lambda}\right)^2 + 2\beta Y_1 + F - \frac{\alpha^2}{\lambda}$$

La nature exacte de Γ varie selon que $\beta = 0$ ou $\beta \neq 0$:

- Si $\beta = 0$, l'équation de Γ s'écrit

$$M(X_1, Y_1) \in \Gamma \Leftrightarrow \left(X_1 + \frac{\alpha}{\lambda}\right)^2 = \frac{\alpha^2}{\lambda^2} - \frac{F}{\lambda}$$

et donc Γ **est vide, formé d'une droite ou de deux droites parallèles** selon que la quantité $\delta = \dfrac{\alpha^2}{\lambda^2} - \dfrac{F}{\lambda}$ est négative, nulle ou positive.

- Si $\beta \neq 0$, l'équation de Γ peut s'écrire

$$\lambda\left(X_1+\frac{\alpha}{\lambda}\right)^2=-2\beta\left(Y_1+\frac{F}{2\beta}-\frac{\alpha^2}{2\lambda\beta}\right)$$

En prenant pour nouvelle origine le point Σ dont les coordonnées, dans le repère $\left(O,\overrightarrow{I}_1,\overrightarrow{J}_1\right)$, sont $\left(-\frac{\alpha}{\lambda},\frac{\alpha^2}{2\lambda\beta}-\frac{F}{2\beta}\right)$, on obtiendra l'équation réduite de Γ dans le repère $\left(\Sigma,\overrightarrow{I}_1,\overrightarrow{J}_1\right)$

$$M(X,Y)\in\Gamma\Leftrightarrow\lambda X^2=-2\beta Y$$

La courbe Γ est donc une parabole de sommet Σ et d'axe parallèle au vecteur $\overrightarrow{J}$: ici encore, **la direction asymptotique** pour Γ correspond à la **droite vectorielle qui annule la forme quadratique Φ.**

20-1.7.3 Centre de symétrie

Dans la présentation générale des calculs qui précède, nous avons d'abord opéré une rotation des axes, puis une translation de l'origine du repère. Dans la pratique, on opère souvent dans l'ordre inverse : lorsque la forme Φ est non dégénérée ($AC-B^2\neq 0$), on commence par chercher un centre de symétrie pour Γ. Comme $Q(x,y)$ est un polynôme du second degré, nous avons, avec les notations de la section **20-1.7.1**

$$\begin{aligned}Q(x_0+X,y_0+Y)=Q(x_0,y_0)&+X\frac{\partial Q}{\partial x}(x_0,y_0)\\&+Y\frac{\partial Q}{\partial y}(x_0,y_0)+AX^2+2BXY+CY^2\end{aligned}$$

Dans l'étude de la section précédente, nous avons vu que nous pouvions trouver un point $\Omega(x_0,y_0)$ tel que cette équation soit formellement invariante en changeant X et Y en $-X$ et $-Y$. On trouvera donc les coordonnées de Ω en résolvant le système

$$\begin{cases}\dfrac{\partial Q}{\partial x}(x_0,y_0)=0\\[2mm]\dfrac{\partial Q}{\partial y}(x_0,y_0)=0\end{cases}$$

Il possède une solution unique : si Ω existe, l'équation

$$\Phi\left(\overrightarrow{OM}\right)+2l\left(\overrightarrow{OM}\right)+F=0$$

s'écrit simplement

$$\Phi\left(\overrightarrow{\Omega M}\right)+F'=0$$

Le calcul fait à la section **20-1.7.1** montre que cela se traduit par

$$\varphi\left(\overrightarrow{O\Omega}, \bullet\right) = -l$$

(φ est la forme polaire de Φ). Comme Φ est non dégénérée, il existe un unique vecteur $\overrightarrow{v}$ vérifiant $\varphi(\overrightarrow{v}, \bullet) = -l$, ce qui prouve l'unicité (et l'existence) de Ω. Dans le repère $\left(\Omega, \overrightarrow{i}, \overrightarrow{j}\right)$, l'équation de Γ s'écrit donc

$$AX^2 + 2BXY + CY^2 + Q(\Omega) = 0$$

Il suffit ensuite de diagonaliser la matrice $\begin{pmatrix} A & B \\ B & C \end{pmatrix}$ pour déterminer les axes de Γ et son équation réduite.

EXERCICE 20-1.42 Décrire la courbe Γ d'équation, dans un repère orthonormé du plan $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$

$$Q(x, y) = x^2 + 4xy + y^2 - 4x - 4y + 2 = 0$$

La matrice M de la forme quadratique est

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Son déterminant vaut -3 et ses valeurs propres sont 3 et -1. Γ est donc du genre hyperbole, et est éventuellement dégénérée en un couple de droites concourantes. Pour déterminer le centre de symétrie $\Omega(x_0, y_0)$, on résout le système

$$\begin{cases} 2x + 4y - 4 = 0 \\ 4x + 2y - 4 = 0 \end{cases}$$

La solution est $(x_0, y_0) = \left(\frac{2}{3}, \frac{2}{3}\right)$. Dans le repère $\left(\Omega, \overrightarrow{i}, \overrightarrow{j}\right)$, l'équation de Γ est donc de la forme $X^2 + 4XY + Y^2 + Q\left(\frac{2}{3}, \frac{2}{3}\right) = 0$, soit

$$X^2 + 4XY + Y^2 = \frac{2}{3}$$

Une base orthonormale diagonalisant l'endomorphisme symétrique associé à M est

$$\overrightarrow{i}_1 = \frac{1}{\sqrt{2}}\left(\overrightarrow{i} + \overrightarrow{j}\right) \quad \text{et} \quad \overrightarrow{j}_1 = \frac{1}{\sqrt{2}}\left(-\overrightarrow{i} + \overrightarrow{j}\right)$$

Les axes de Γ sont donc parallèles aux bissectrices des axes du repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$, et son équation réduite dans le repère $\left(\Omega, \overrightarrow{i}_1, \overrightarrow{j}_1\right)$ est

$$3X_1^2 - Y_1^2 = \frac{2}{3} \Leftrightarrow \frac{X_1^2}{\frac{2}{9}} - \frac{Y_1^2}{\frac{2}{3}} = 1$$

soit $a^2 = \frac{2}{9}$ et $b^2 = \frac{2}{3}$. L'axe focal est donc $\Omega + \operatorname{vect}\left(\overrightarrow{i}\right)$, c'est la première bissectrice du repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$.

REMARQUE 20-1.43 Dans le cas d'une courbe du genre hyperbole, on peut trouver aussi les coordonnées du centre et les équations des asymptotes par la méthode suivante :

On sait que l'équation d'une hyperbole dans un repère dont les axes sont les asymptotes est de la forme

$$XY = k$$

(le cas $k = 0$ correspondant au cas de deux droites concourantes). Lorsque Φ est de signature $(1,1)$, elle est différence de deux carrés donc aussi produit de deux formes linéaires indépendantes,

$$\Phi = l_1^2 - l_2^2 = l_1' l_2' \quad \text{avec} \quad l_1' = l_1 + l_2 \text{ et } l_2' = l_1 - l_2$$

Si l'équation de Γ s'écrit comme d'habitude

$$\Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + F = 0$$

la forme linéaire $2l$ peut s'écrire comme combinaison linéaire de l_1' et l_2' (qui forment une base du dual de l'espace vectoriel E_2)

$$2l = \alpha\, l_1' + \beta\, l_2'$$

avec $(\alpha, \beta) \in \mathbb{R}^2$. L'équation de Γ fait alors intervenir le produit de deux formes affines

$$\Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) = -F \Leftrightarrow \left(l_1'\left(\overrightarrow{OM}\right) + \beta\right)\left(l_2'\left(\overrightarrow{OM}\right) + \alpha\right) = \alpha\beta - F$$

et on obtient ainsi les équations des asymptotes à Γ

$$l_1'\left(\overrightarrow{OM}\right) + \beta = 0 \quad \text{et} \quad l_2'\left(\overrightarrow{OM}\right) + \alpha = 0$$

Par exemple, pour la courbe Γ étudiée plus haut, d'équation

$$x^2 + 4xy + y^2 - 4x - 4y + 2 = 0$$

on écrit

$$x^2 + 4xy + y^2 = (x + 2y)^2 - 3y^2 = \left(x + \left(2 + \sqrt{3}\right)y\right)\left(x + \left(2 - \sqrt{3}\right)y\right)$$

En posant

$$l_1'(x,y) = x + \left(2 + \sqrt{3}\right)y \text{ et } l_2'(x,y) = x + \left(2 - \sqrt{3}\right)y$$

on a

$$y = \frac{\sqrt{3}}{6}\left(l_1'(x,y) - l_2'(x,y)\right)$$

ce qui donne

$$\begin{aligned} 4(x+y) &= 4l_1'(x,y) - 4\left(1 + \sqrt{3}\right)y \\ &= 4l_1'(x,y) - 4\left(1 + \sqrt{3}\right)\left(\frac{\sqrt{3}}{6}\left(l_1'(x,y) - l_2'(x,y)\right)\right) \end{aligned}$$

L'équation de Γ s'écrit donc

$$l'_1 l'_2 - \left(2 - \frac{2\sqrt{3}}{3}\right) l'_1 - \left(2 + \frac{2\sqrt{3}}{3}\right) l'_2 + 2 = 0$$

soit

$$\left(l'_1 - \frac{6 + 2\sqrt{3}}{3}\right)\left(l'_2 - \frac{6 - 2\sqrt{3}}{3}\right) = \frac{2}{3}$$

Les équations des asymptotes de Γ sont donc

$$x + \left(2 + \sqrt{3}\right) y - \frac{6 + 2\sqrt{3}}{3} = 0 \quad \text{et} \quad x + \left(2 - \sqrt{3}\right) y - \frac{6 - 2\sqrt{3}}{3} = 0$$

On vérifie facilement que l'intersection de ces deux droites est le centre $\Omega\left(\frac{2}{3}, \frac{2}{3}\right)$ de Γ.

20-1.7.4 Coniques : définition par foyer et directrice

Dans toute cette section, F est un point et $\mathcal{D}$ est une droite du plan euclidien $\mathcal{E}_2$ avec $F \notin \mathcal{D}$.

DÉFINITION 20-1.44 *Soit e un réel > 0. On appelle conique de foyer F, de directrice $\mathcal{D}$ et d'excentricité e l'ensemble des points du plan dont le rapport des distances à F et $\mathcal{D}$ est égal à e :*

$$\mathcal{C} = \left\{ M \in \mathcal{E}_2 \mid \frac{d(M,F)}{d(M,\mathcal{D})} = e \right\}$$

Soit H_0 la projection orthogonale de F sur $\mathcal{D}$, avec $\left\|\overrightarrow{FH_0}\right\| = h > 0$. Choisissons un repère orthonormé $\left(F, \overrightarrow{I}, \overrightarrow{J}\right)$ du plan, d'origine F, et tel que l'axe des abscisses soit la droite FH_0, avec donc $H_0(h,0)$. Si $M(x,y) \in \mathcal{E}_2$, sa projection orthogonale H sur $\mathcal{D}$ a pour coordonnées (h,y) et

$$d(M,\mathcal{D}) = MH = |x - h|$$

On a donc

$$M \in \mathcal{C} \Leftrightarrow MF^2 = e^2 MH^2$$

ce qui donne la condition

$$M(x,y) \in \mathcal{C} \Leftrightarrow x^2 + y^2 - e^2 (x-h)^2 = 0$$

Il s'agit donc d'une courbe du second degré, associé à la forme quadratique définie sur E_2 par

$$\Phi\left(x\overrightarrow{I} + y\overrightarrow{J}\right) = \left(1 - e^2\right) x^2 + y^2$$

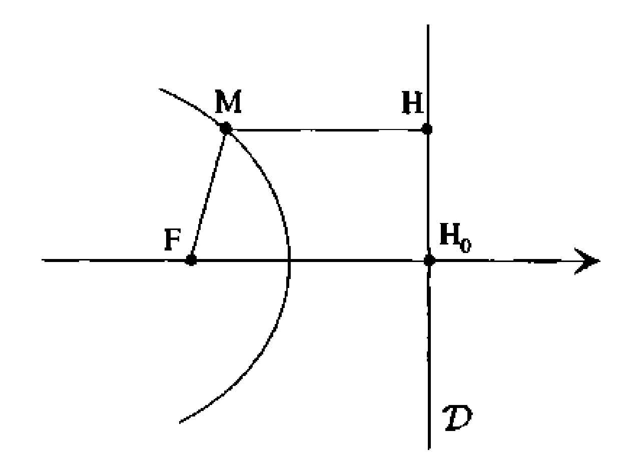


Figure **20.11** – *Définition par foyer et directrice :* $\dfrac{MF}{MH} = e$

Elle est donc du genre :

$$\mathcal{C} : \begin{cases} \bullet \ \textbf{ELLIPSE} \quad \text{si} \quad e < 1 \\ \bullet \ \textbf{PARABOLE} \quad \text{si} \quad e = 1 \\ \bullet \ \textbf{HYPERBOLE} \quad \text{si} \quad e > 1 \end{cases}$$

- Lorsque $e = 1$, l'équation s'écrit $y^2 = -2hx + h^2 = -2h\left(x - \dfrac{h}{2}\right)$, et le sommet de la parabole est $S\left(\dfrac{h}{2}, 0\right)$. En ramenant l'origine au sommet, on obtient

PROPOSITION 20-1.45 Dans un repère orthonormé, le foyer de la parabole d'équation $y^2 = 2px$ est le point $F\left(\frac{p}{2}, 0\right)$ et la directrice a pour équation $x = -\frac{p}{2}$ (le réel $|p|$ est appelé paramètre de la parabole. C'est la longueur de la demi-corde passant par le foyer et parallèle à la directrice).

- Lorsque $e \neq 1$, on peut chercher le centre de symétrie, dont l'abscisse vérifie

$$\frac{\partial}{\partial x}\left(x^2 + y^2 - e^2(x-h)^2\right) = 2\left(\left(1-e^2\right)x + e^2h\right) = 0$$

ce qui donne

$$S\left(\frac{e^2h}{e^2-1}, 0\right)$$

Dans le repère $\left(S, \vec{I}, \vec{J}\right)$, l'équation de $\mathcal{C}$ est donc

$$\left(1-e^2\right)X^2 + Y^2 + Q(S) = 0$$

(notation de la section précédente), soit

$$\left(1-e^2\right)X^2 + Y^2 = \frac{e^2h^2}{1-e^2}$$

- Si $e < 1$, on a bien l'équation d'une ellipse de centre S, d'équation

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$$

avec

$$a^2 = \frac{e^2h^2}{(1-e^2)^2} > b^2 = \frac{e^2h^2}{(1-e^2)}$$

Dans ce repère, le foyer F a pour coordonnées $(c, 0)$ avec

$$c = \frac{e^2h}{1-e^2}$$

et la directrice a pour équation $X = X_{H_0}$ avec

$$X_{H_0} = h - \frac{e^2h}{e^2-1} = \frac{h}{e^2-1}$$

On vérifie aisément que $a^2 - b^2 = c^2$, $e = \frac{c}{a}$ et $X_{H_0} = \frac{a^2}{c}$. Enfin, par définition de $\mathcal{C}$, la longueur de la demi-corde focale parallèle à la directrice est $p = eh = \frac{b^2}{a}$. On obtient donc :

PROPOSITION 20-1.46 Dans un repère orthonormé, l'ellipse d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

avec $a > b > 0$ est une conique de foyer $F\left(c = \sqrt{a^2-b^2}, 0\right)$, de directrice d'équation $x = \frac{a^2}{c}$ et d'excentricité $e = \frac{c}{a} < 1$. Le paramètre vaut $p = \frac{b^2}{a}$. Pour des raisons de symétrie, il existe un autre couple de foyer et directrice. Ce sont $F'(-c, 0)$ et la droite $\mathcal{D}'$ d'équation $x = -\frac{a^2}{c}$.

- Si $e > 1$, $\mathcal{C}$ est une hyperbole. En menant les calculs comme dans le cas de l'ellipse, on obtient :

PROPOSITION 20-1.47 Dans un repère orthonormé, l'hyperbole d'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

est une conique de foyer $F\left(c = \sqrt{a^2+b^2}, 0\right)$, de directrice d'équation $x = \frac{a^2}{c}$ et d'excentricité $e = \frac{c}{a} > 1$. Le paramètre vaut $p = \frac{b^2}{a}$. $F'(-c, 0)$ et la droite $\mathcal{D}'$ d'équation $x = -\frac{a^2}{c}$ sont un autre couple de foyer et directrice. Les asymptotes ont pour équations

$$y = \pm\frac{b}{a}x$$

Ces asymptotes sont orthogonales (hyperbole équilatère) si et seulement si $a = b$, ce qui équivaut à $e = \sqrt{2}$.

20-1.7.5 Définition bifocale des coniques à centre

EXERCICE 20-1.48 Soient F et F' deux points du plan euclidien avec $FF' = 2c > 0$. Si a est un réel strictement supérieur à c, montrer que

$$\Gamma = \{M \in \mathcal{E}_2 \mid MF + MF' = 2a\}$$

est une ellipse de foyers F et F' et de demi grand axe égal à a.

La condition $a > c$ est évidemment imposée par l'inégalité triangulaire (le cas $a = c$ étant inintéressant). On choisit un repère $\left(O, \overrightarrow{i}, \overrightarrow{j}\right)$ orthonormé tel que F et F' aient pour coordonnées $(\pm c, 0)$. Soit $M \in \mathcal{E}_2$ de coordonnées (x, y). On a

$$MF = \sqrt{(x-c)^2 + y^2} \quad \text{et} \quad MF' = \sqrt{(x+c)^2 + y^2}$$

ce qui donne évidemment

$$MF'^2 - MF^2 = 4xc$$

Si $M \in \Gamma$, on a

$$4cx = \left(MF' - MF\right) 2a$$

et donc

$$MF' - MF = \frac{2cx}{a} \text{ et } MF + MF' = 2a \Rightarrow MF = a - \frac{cx}{a}$$

On en déduit

$$MF^2 = (x-c)^2 + y^2 = \left(a - \frac{cx}{a}\right)^2$$

En posant $b^2 = a^2 - c^2$, cette condition s'écrit

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

elle est nécessairement vérifiée si $M \in \Gamma$. Réciproquement, si $M(x, y)$ vérifie cette équation, on a $|x| \leqslant a$, et donc $a - \frac{cx}{a} > 0$. En reprenant les calculs en sens inverse, on obtient $MF = a - \frac{cx}{a}$, ce qui donne $MF + MF' = 2a$, compte tenu de $MF' - MF = \frac{2cx}{a}$.

EXERCICE 20-1.49 Montrer de même que, si on prend $0 < a < c$, l'ensemble des points du plan vérifiant la condition

$$\left|MF - MF'\right| = 2a$$

est un hyperbole de foyers F et F', et d'excentricité $\frac{c}{a}$.

EXERCICE 20-1.50 Montrer que la normale en tout point M d'une ellipse de foyers F et F' est bissectrice intérieure de MF et MF'. (Indication : le plus simple est d'utiliser un paramétrage régulier quelconque $t \mapsto M(t)$ et d'utiliser le fait que l'application $t \mapsto \left\|\overrightarrow{FM(t)}\right\| + \left\|\overrightarrow{F'M(t)}\right\|$ est constante). Montrer que cette propriété caractérise les arcs réguliers dont le support est inclus dans une ellipse de foyers F et F'. Enoncer un résultat analogue pour les hyperboles de foyers F et F'.

20-1.7.6 Equation polaire des coniques (origine au foyer)

Dans le repère $\left(F, \vec{\imath}, \vec{\jmath}\right)$ utilisé au début de la section **20-1.7.4**, où la droite $\mathcal{D}$ a pour équation $x = h$, la conique $\mathcal{C}$ de foyer F, de directrice $\mathcal{D}$ et d'excentricité e a pour équation

$$x^2 + y^2 = e^2 (x - h)^2$$

Si (r, θ) est un système de coordonnées polaires d'un point M dans $\left(F, \vec{\imath}, \vec{\jmath}\right)$, on a donc

$$M \in \mathcal{C} \Leftrightarrow r^2 = e^2 (r \cos \theta - h)^2$$

ce qui équivaut à

$$r = er \cos \theta - eh \quad \text{ou} \quad r = -er \cos \theta + eh$$

Si le couple (ρ, α) vérifie la première équation, le couple $(-\rho, \alpha + \pi)$ vérifie la seconde et représente le même point du plan. Il en résulte que l'équation polaire

$$r = \frac{eh}{1 + e \cos \theta} = \frac{p}{1 + e \cos \theta}$$

représente la totalité de $\mathcal{C}$.

PROPOSITION 20-1.51 L'arc défini dans le repère $\left(O, \vec{\imath}, \vec{\jmath}\right)$ par l'équation polaire

$$r = \frac{p}{1 + e \cos \theta}$$

est la conique de foyer O, d'excentricité e et de directrice $\mathcal{D}$ d'équation $x = \frac{p}{e}$. La directrice $\mathcal{D}$ a donc pour équation polaire

$$\mathcal{D} : r = \frac{p}{e \cos \theta}$$

Si on fait tourner la directrice d'un angle α, l'équation de la conique devient

$$r = \frac{p}{1 + e \cos (\theta - \alpha)}$$

et la directrice a alors pour équation

$$r = \frac{p}{e \cos (\theta - \alpha)}$$

En utilisant les formules de transformation trigonométrique, on obtient facilement :

COROLLAIRE 20-1.52 L'équation polaire générale d'une conique de foyer O (origine du repère) est de la forme

$$r = \frac{1}{a + b\cos\theta + c\sin\theta}$$

avec $a \neq 0$ et $(b,c) \neq (0,0)$. L'excentricité est alors

$$e = \sqrt{\frac{b^2+c^2}{a^2}}$$

et la directrice a pour équation polaire

$$r = \frac{1}{b\cos\theta + c\sin\theta}$$

EXEMPLE 20-1.53 Reprenons la courbe étudiée à l'exercice **20-1**.42 : son équation dans $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$ est

$$(E) : x^2 + 4xy + y^2 - 4x - 4y + 2 = 0$$

Il s'agit d'une hyperbole de centre $\Omega\left(\frac{2}{3}, \frac{2}{3}\right)$ et d'équation réduite

$$\frac{X_1^2}{\frac{2}{9}} - \frac{Y_1^2}{\frac{2}{3}} = 1$$

dans le repère $\left(\Omega, \overrightarrow{I}_1, \overrightarrow{J}_1\right)$ avec

$$\overrightarrow{I}_1 = \frac{1}{\sqrt{2}}\left(\overrightarrow{I} + \overrightarrow{J}\right) \quad \text{et} \quad \overrightarrow{J}_1 = \frac{1}{\sqrt{2}}\left(-\overrightarrow{I} + \overrightarrow{J}\right)$$

On a donc

$$c^2 = \frac{2}{9} + \frac{2}{3} \Rightarrow c = \frac{2\sqrt{2}}{3}$$

Par conséquent, un des foyers F vérifie

$$\overrightarrow{\Omega F} = -c\overrightarrow{I}_1 = -\frac{2}{3}\left(\overrightarrow{I} + \overrightarrow{J}\right) = \overrightarrow{\Omega O}$$

L'origine O est donc un des foyers de l'hyperbole. Nous pouvons obtenir une équation polaire de celle-ci à partir de l'équation (E) : le couple $(r\cos\theta, r\sin\theta)$ vérifie (E) si et seulement si

$$r^2(1 + 4\cos\theta\sin\theta) - 4r(\cos\theta + \sin\theta) + 2 = 0$$

Pour $\theta \neq -\dfrac{\pi}{3} \pmod{2\pi}$, on obtient une équation du second degré en r. Son discriminant réduit est $\Delta' = 4(\cos\theta + \sin\theta)^2 - 2(1 + 4\cos\theta\sin\theta) = 2$ et ses racines sont donc

$$r_1 = \frac{2(\cos\theta + \sin\theta) + \sqrt{2}}{(1 + 4\cos\theta\sin\theta)} \quad \text{et} \quad r_2 = \frac{2(\cos\theta + \sin\theta) - \sqrt{2}}{(1 + 4\cos\theta\sin\theta)}$$

Le produit des racines de cette équation vaut $\dfrac{2}{1+4\cos\theta\sin\theta}$, ce qui donne

$$r_1 = \frac{2}{2(\cos\theta+\sin\theta)-\sqrt{2}} \quad \text{et} \quad r_2 = \frac{2}{2(\cos\theta+\sin\theta)+\sqrt{2}}$$

Ici encore, le couple (r,θ) vérifie la première égalité si et seulement si $(-r,\theta+\pi)$ vérifie la seconde. Nous garderons donc l'équation

$$r = \frac{2}{2(\cos\theta+\sin\theta)+\sqrt{2}} = \frac{\sqrt{2}}{1+2\cos\left(\theta-\frac{\pi}{4}\right)}$$

Cette équation est bien de la forme voulue. On retrouve le fait que l'axe contenant les foyers fait un angle de $\dfrac{\pi}{4}$ avec l'axe des abscisses. L'excentricité est 2, et la directrice $\mathcal{D}$ a pour équation

$$r = \frac{1}{\cos\theta+\sin\theta} \quad \text{soit} \quad x+y+1=0$$

On vérifie effectivement que l'équation de départ peut s'écrire

$$x^2+y^2 = 4\left(\frac{x+y+1}{\sqrt{2}}\right)^2$$

ce qui traduit bien la condition

$$d^2(M,O) = 4d^2(M,\mathcal{D})$$

EXERCICE 20-1.54 On considère la conique de foyer O et d'équation polaire

$$r = \frac{p}{1+e\cos\theta}$$

Montrer qu'une équation polaire de la tangente au point de paramètre θ_0 est

$$\frac{1}{r} = \frac{1+e\cos\theta_0}{p}\cos(\theta-\theta_0) - \frac{e\sin\theta_0}{p}\sin(\theta-\theta_0)$$

En déduire que la portion de tangente en $M(\theta_0)$ comprise entre $M(\theta_0)$ et la directrice est vue du foyer sous un angle droit.

EXERCICE 20-1.55 Etudier également le nombre de tangentes qu'on peut mener d'un point P de coordonnées polaires (ρ,ϕ) à cette conique. Lorsqu'on peut mener deux tangentes $\mathcal{T}_1$ et $\mathcal{T}_2$ rencontrant la conique respectivement en M_1 et M_2, montrer que OP est bissectrice de OM_1 et OM_2.

EXERCICE 20-1.56 Déterminer le lieu des centres des hyperboles équilatères de foyer O passant par un point fixe A.

20-2 Etude métrique

20-2.1 Longueur d'un arc

20-2.1.1 Arc rectifiable

Nous considérons dans cette section un arc de classe C^k ($k \geqslant 0$)

$$\Gamma : [a,b] \ni t \mapsto M(t) \in \mathcal{E}_n$$

à valeurs dans un espace affine euclidien de dimension n. Si

$$d : a = t_0 < t_1 < \cdots < t_p = b$$

est une subdivision quelconque de $[a,b]$ (on suppose $a < b$), on peut lui associer la famille de points $(M_i = M(t_i))_{0 \leqslant i \leqslant p}$ et la ligne polygonale

$$\bigcup_{i=0}^{p-1} [M_i M_{i+1}]$$

inscrite dans le support de l'arc. On note

$$L(d) = \sum_{i=0}^{p-1} \left\| \overrightarrow{M_i M}_{i+1} \right\|$$

la longueur de cette ligne polygonale.

DÉFINITION 20-2.1 *On dit que l'arc Γ est rectifiable si et seulement si le sous-ensemble de $\mathbb{R}^+$*

$$\{L(d) \mid d \text{ subdivision de } [a,b]\}$$

est majoré. Dans ce cas, sa borne supérieure est appelée longueur de l'arc Γ. On notera

$$L(\Gamma) = \sup \{L(d) \mid d \text{ subdivision de } [a,b]\}$$

Attention ! L'interprétation physique de cette notion est la longueur du chemin parcouru par le point mobile $M(t)$ pour t décrivant $[a,b]$. Il ne faut pas confondre avec la longueur du support de Γ, même si les deux notions se rejoignent lorsque l'arc est simple. En tout cas, il s'agit d'une notion géométrique :

PROPOSITION 20-2.2 Si $\Gamma_1 : [a,b] \ni t \mapsto M_1(t)$ et $\Gamma_2 : [\alpha,\beta] \ni s \mapsto M_2(s)$ sont deux arcs C^k-équivalents, Γ_1 est rectifiable si et seulement si Γ_2 l'est, et on a alors

$$L(\Gamma_1) = L(\Gamma_2)$$

Démonstration : Il suffit de revenir à la définition de l'équivalence.

■

Si Γ représente l'arc $[a,b] \ni t \mapsto M(t)$ et si $[a',b'] \subset [a,b]$, nous noterons $\Gamma|_{[a',b']}$ l'arc obtenu par restriction au segment $[a',b']$.

PROPOSITION 20-2.3 Si $\Gamma : [a,b] \ni t \mapsto M(t)$ est rectifiable, il en est de même de $\Gamma|_{[a',b']}$ pour $[a',b'] \subset [a,b]$, et on a

$$L\left(\Gamma|_{[a',b']}\right) \leqslant L(\Gamma)$$

Démonstration : Supposons $a < a' < b' < b$ (les cas particuliers $a = a'$ ou $b = b'$ se traitent de manière analogue). Si d' est une subdivision de $[a',b']$, on obtient une subdivision de $[a,b]$ en rajoutant les points a et b, et on a alors

$$\left\|\overrightarrow{M(a)\,M(a')}\right\| + L(d') + \left\|\overrightarrow{M(b')\,M(b)}\right\| = L(d) \leqslant L(\Gamma)$$

ce qui donne en particulier $L(d') \leqslant L(\Gamma)$. On en déduit que l'arc $\Gamma' = \Gamma|_{[a',b']}$ est rectifiable, avec $L(\Gamma') \leqslant L(\Gamma)$. ■

COROLLAIRE 20-2.4 Soit $\Gamma : [a,b] \ni t \mapsto M(t)$ un arc de classe C^k ($k \geqslant 0$) et $c \in]a,b[$ quelconque. Γ est rectifiable si et seulement si $\Gamma|_{[a,c]}$ et $\Gamma|_{[c,b]}$ le sont, et on a alors

$$L(\Gamma) = L\left(\Gamma|_{[a,c]}\right) + L\left(\Gamma|_{[c,b]}\right)$$

Démonstration : Si Γ est rectifiable alors $\Gamma|_{[a,c]}$ et $\Gamma|_{[c,b]}$ le sont, d'après la proposition précédente. Montrons la réciproque : si d une subdivision de $[a,b]$, en rajoutant éventuellement le point c, on obtient une subdivision d_1 de $[a,b]$, qu'on peut interpréter comme juxtaposition d'une subdivision d' de $[a,c]$ et d'une subdivision d'' de $[c,b]$. On a clairement, à cause de l'inégalité triangulaire,

$$L(d) \leqslant L(d_1) = L(d') + L(d'')$$

Si on suppose $\Gamma|_{[a,c]}$ et $\Gamma|_{[c,b]}$ rectifiables, on obtient

$$L(d) \leqslant L\left(\Gamma|_{[a,c]}\right) + L\left(\Gamma|_{[c,b]}\right)$$

ce qui montre que Γ est bien rectifiable avec

$$L(\Gamma) \leqslant L\left(\Gamma|_{[a,c]}\right) + L\left(\Gamma|_{[c,b]}\right)$$

L'inégalité inverse se prouve par la même méthode : si d' et d'' sont des subdivisions respectives de $[a,c]$ et $[c,b]$, la juxtaposition de d' et d'' donne une subdivision de $[a,b]$, avec

$$L(d') + L(d'') = L(d) \leqslant L(\Gamma)$$

ce qui donne en particulier

$$L(d') \leqslant L(\Gamma) - L(d'')$$

Comme cette inégalité est valable quelle que soit d', on en déduit

$$L\left(\Gamma|_{[a,c]}\right) \leqslant L(\Gamma) - L(d'')$$

soit encore $L(d'') \leqslant L(\Gamma) - L\left(\Gamma|_{[a,c]}\right)$. On obtient finalement, puisque la majoration est valable avec d'' quelconque,

$$L\left(\Gamma|_{[c,b]}\right) \leqslant L(\Gamma) - L\left(\Gamma|_{[a,c]}\right)$$

On a donc bien

$$L(\Gamma) = L\left(\Gamma|_{[a,c]}\right) + L\left(\Gamma|_{[c,b]}\right) \quad \blacksquare$$

Cette égalité est connue sous le nom de relation de Chasles. Pour qu'elle soit encore valable pour $c = a$ ou $c = b$, on attribue par convention une longueur nulle à un "arc" dont le segment de définition est réduit à un point.

REMARQUE 20-2.5 Si $\Gamma : [a,b] \ni t \mapsto M(t)$ est un arc rectifiable, le choix de la subdivision

$$d : a = t_0 < t_1 = b$$

donne immédiatement

$$\left\|\overrightarrow{M(a)\,M(b)}\right\| \leqslant L(\Gamma)$$

La longueur de la corde $[M(a), M(b)]$ est inférieure à celle de l'arc.

20-2.1.2 Longueur d'un arc de classe C^1

THÉORÈME 20-2.6 Tout arc $\Gamma : [a,b] \ni t \mapsto M(t)$ de classe C^k ($k \geqslant 1$) est rectifiable avec

$$L(\Gamma) = \int_a^b \left\|\overrightarrow{M}'(t)\right\| dt$$

Démonstration : soit $d : a = t_0 < t_1 < \cdots < t_p = b$ une subdivision de $[a,b]$. Si on note comme précédemment $M_i = M(t_i)$, on a, puisque l'arc est de classe C^1,

$$\forall i \quad \overrightarrow{M_i M}_{i+1} = \int_{t_i}^{t_{i+1}} \overrightarrow{M}'(t)\, dt$$

ce qui donne, par inégalité de la norme,

$$\left\|\overrightarrow{M_i M}_{i+1}\right\| \leqslant \int_{t_i}^{t_{i+1}} \left\|\overrightarrow{M}'(t)\right\| dt$$

et finalement

$$L(d) = \sum_{i=0}^{p-1} \left\|\overrightarrow{M_i M}_{i+1}\right\| \leqslant \sum_{i=0}^{p-1} \int_{t_i}^{t_{i+1}} \left\|\overrightarrow{M}'(t)\right\| dt = \int_a^b \left\|\overrightarrow{M}'(t)\right\| dt$$

L'arc Γ est donc rectifiable et sa longueur vérifie

$$L(\Gamma) \leqslant \int_a^b \left\| \overrightarrow{M}'(t) \right\| dt$$

Pour prouver l'égalité, nous montrons que l'application

$$\varphi : [a,b] \to \mathbb{R}^+ \quad x \mapsto \varphi(x) = L\left(\Gamma|_{[a,x]}\right)$$

(bien définie, puisque $\Gamma|_{[a,x]}$ est rectifiable) est de classe C^1, avec $\varphi'(x) = \left\| \overrightarrow{M}'(x) \right\|$. Nous prouverons la dérivabilité à droite en tout point de $[a,b[$, la dérivabilité à gauche se prouvant de la même façon. Soit $x \in [a,b[$ et $h > 0$ avec $x + h \in [a,b]$. Nous avons par la relation de Chasles

$$\varphi(x+h) - \varphi(x) = L\left(\Gamma|_{[x,x+h]}\right)$$

Comme $\Gamma|_{[x,x+h]}$ est de classe C^1, nous avons, d'après ce qui a été démontré plus haut

$$L\left(\Gamma|_{[x,x+h]}\right) \leqslant \int_x^{x+h} \left\| \overrightarrow{M}'(t) \right\| dt$$

Comme la longueur de la corde est inférieure à celle de l'arc, on a

$$\left\| \overrightarrow{M(x)\,M(x+h)} \right\| \leqslant L\left(\Gamma|_{[x,x+h]}\right)$$

On obtient donc

$$\left\| \frac{\overrightarrow{M(x)\,M(x+h)}}{h} \right\| \leqslant \frac{\varphi(x+h) - \varphi(x)}{h} \leqslant \frac{1}{h} \int_x^{x+h} \left\| \overrightarrow{M}'(t) \right\| dt$$

Le théorème de dérivation d'une intégrale fonction d'une de ses bornes montrant que

$$\lim_{h \to 0} \frac{1}{h} \int_x^{x+h} \left\| \overrightarrow{M}'(t) \right\| dt = \left\| \overrightarrow{M}'(x) \right\|$$

l'encadrement précédent donne

$$\lim_{h \to 0^+} \frac{\varphi(x+h) - \varphi(x)}{h} = \left\| \overrightarrow{M}'(x) \right\|$$

La fonction φ est bien de classe C^1, et on a

$$L(\Gamma) = \varphi(b) - \varphi(a) = \int_a^b \varphi'(t)\, dt = \int_a^b \left\| \overrightarrow{M}'(t) \right\| dt \quad \blacksquare$$

COROLLAIRE 20-2.7 Tout arc $\Gamma : [a,b] \ni t \mapsto M(t)$ continu et C^1 par morceaux est rectifiable avec

$$L(\Gamma) = \int_a^b \left\| \overrightarrow{M}'(t) \right\| dt$$

(intégrale d'une fonction continue par morceaux si on la prolonge arbitrairement en tout point de non dérivabilité de M)

Démonstration : Il suffit d'appliquer la relation de Chasles aux deux membres de l'égalité, en utilisant une subdivision de $[a,b]$ adaptée à l'application C^1 par morceaux $t \mapsto M(t)$. ■

REMARQUE 20-2.8 Un arc supposé seulement continu n'est pas nécessairement rectifiable. On peut par exemple considérer l'arc tracé dans $\mathbb{R}^2$ euclidien défini par

$$[0,1] \ni t \mapsto M(t) = (t, \varphi(t))$$

avec $\varphi(0) = 0$ et, pour $n \in \mathbb{N}^*$ arbitraire,

$$\varphi(t) = (-1)^n \left[\left((n+1)\sqrt{n} + n\sqrt{n+1} \right) \left(t - \frac{1}{n+1} \right) - \frac{1}{\sqrt{n+1}} \right]$$

si $t \in \left] \frac{1}{n+1}, \frac{1}{n} \right]$ Représenter le support de cet arc, montrer qu'il est continu et dire pourquoi il n'est pas rectifiable.

REMARQUE 20-2.9 Nous n'avons utilisé nulle part le fait que la norme avec laquelle on travaille est associée à un produit scalaire. Cette hypothèse interviendra ultérieurement pour définir la courbure. L'étude précédente permet en fait de rectifier des arcs dans un espace affine basé sur un espace normé complet quelconque.

20-2.1.3 Abscisse curviligne

DÉFINITION 20-2.10 *Soit $\Gamma : I \ni t \mapsto M(t)$ un arc de classe C^1 défini sur un intervalle I de $\mathbb{R}$ et $t_0 \in I$ choisi arbitrairement. On appelle abscisse curviligne de Γ pour l'origine t_0 l'application*

$$s : I \to \mathbb{R} \qquad t \mapsto s(t) = \begin{cases} L\left(\Gamma|_{[t_0,t]}\right) & si \quad t \geqslant t_0 \\ -L\left(\Gamma|_{[t,t_0]}\right) & si \quad t \leqslant t_0 \end{cases}$$

On a donc, conformément aux résultats de la section précédente,

$$s(t) = \int_{t_0}^{t} \left\| \overrightarrow{M}'(u) \right\| du$$

L'application s est donc aussi de classe C^1. Le choix d'une autre origine dans I modifie cette application s en lui rajoutant une constante. La relation de Chasles donne évidemment

$$\forall a < b \in I \qquad L\left(\Gamma|_{[a,b]}\right) = s(b) - s(a)$$

REMARQUE 20-2.11 On pourrait définir de même l'abscisse curviligne sur un arc continu et C^1 par morceaux.

20-2.2 Courbure des arcs plans

Dans toute cette section, on suppose que $\mathcal{E}_2$ est un **plan euclidien orienté** par le choix d'un repère orthonormé direct $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$.

20-2.2.1 Paramétrage normal d'un arc régulier

DÉFINITION 20-2.12 *Soit* Γ_1 *un arc de classe* C^k $(k \geqslant 1)$

$$J \ni u \mapsto P(u)$$

on dit que Γ_1 *est un paramétrage normal de l'arc géométrique associé si et seulement si*

$$\forall u \in J \quad \left\|\overrightarrow{P}'(u)\right\| = 1$$

L'arc géométrique associé à Γ_1 est donc évidemment régulier, puisque le vecteur dérivé ne s'annule jamais. Réciproquement, nous voyons que tout arc régulier possède des paramétrages admissibles normaux :

PROPOSITION 20-2.13 Soit $\Gamma : I \ni t \mapsto M(t)$ **un arc régulier de classe** C^k **défini sur un intervalle** I **de** $\mathbb{R}$**, et** $t_0 \in I$ **arbitraire. L'application abscisse curviligne pour l'origine** t_0

$$s : I \to \mathbb{R} \quad t \mapsto \int_{t_0}^{t} \left\|\overrightarrow{M}'(u)\right\| du$$

est un C^k**-difféomorphisme de** I **dans** $J = s(I)$ **qu'on peut donc interpréter comme changement de paramétrage admissible. Si on note** φ **le difféomorphisme réciproque**

$$J \ni s \mapsto \varphi(s) = t \in I$$

le paramétrage

$$J \ni s \mapsto P(s) = M(\varphi(s))$$

est un paramétrage normal de l'arc géométrique associé à Γ**.**

Démonstration : L'application s est évidemment de classe C^1, avec

$$\forall t \in I \quad s'(t) = \left\|\overrightarrow{M}'(t)\right\| \neq 0$$

Pour voir que s définit un C^k difféomorphisme de I dans $J = s(I)$, il suffit de voir que s est effectivement de classe C^k. Ceci est conséquence du fait que la norme utilisée est euclidienne : si $(x(t), y(t))$ sont les coordonnées de $M(t)$ dans le repère orthonormé $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$, nous avons

$$\forall t \in I \quad s'(t) = \sqrt{x'^2(t) + y'^2(t)}$$

ce qui prouve que s' est de classe C^{k-1} (parce que la quantité sous le radical ne s'annule jamais). Le paramétrage

$$J \ni s \mapsto P(s) = M(\varphi(s))$$

est alors normal puisque

$$\forall s \in J \quad \overrightarrow{P}'(s) = \varphi'(s)\, \overrightarrow{M}'(\varphi(s)) = \frac{1}{\left\|\overrightarrow{M}'(\varphi(s))\right\|}\overrightarrow{M}'(\varphi(s))$$

d'après la formule de dérivation d'une fonction réciproque. ■

On dit, de manière un peu imprécise, qu'un arc régulier de classe C^k ($k \geqslant 1$) peut être paramétré par son abscisse curviligne. Remarquons que le changement de paramètre envisagé dans cette proposition correspond à un C^k-difféomorphisme strictement croissant. A une translation (sur le paramètre) près, c'est le seul changement de paramètre respectant l'orientation de l'arc dont le résultat soit un paramétrage normal :

PROPOSITION 20-2.14 Soit $\Gamma_1 : J \ni s \mapsto P(s)$ un paramétrage normal d'un arc régulier de classe C^k. Si $K \ni u \mapsto N(u)$ est un arc équivalent dont le paramétrage est également normal, alors le C^k difféomorphisme correspondant

$$\psi : J \to K \quad s \mapsto u = \psi(s)$$

est de la forme $u = \psi(s) = s + \alpha$ ou $u = \psi(s) = \alpha - s$ (où α est une constante), selon que ψ est strictement croissant ou strictement décroissant.

Démonstration : Comme $P = N \circ \psi$, nous avons

$$\forall s \in J \quad \overrightarrow{P}'(s) = \psi'(s)\, \overrightarrow{N}'(\psi(s))$$

d'où $\left\|\overrightarrow{P}'(s)\right\| = |\psi'(s)| \left\|\overrightarrow{N}'(\psi(s))\right\|$. Comme les deux paramétrages sont supposés normaux, nous avons $|\psi'(s)| = 1$ pour tout s de I. Comme ψ' est continue, ceci entraîne que ψ' est constant égal à $+1$ ou -1. Le résultat en découle. ■

NOTATIONS : lorsqu'on se donne un arc de classe C^k $I \ni t \mapsto M(t)$ régulier, on est souvent amené à considérer un paramétrage normal équivalent (essentiellement l'abscisse curviligne). L'usage veut qu'on ne change pas le nom de l'application M, et que l'on note simplement $J \ni s \mapsto M(s)$ ce paramétrage. En particulier, nous nous permettrons des abus d'écriture de la forme

$$M_0 = M(s_0) = M(t_0)$$

lorsque les points t_0 et s_0 se correspondent par le difféomorphisme $I \to J$ associé au changement de paramétrage[9]. De même, nous utiliserons l'écriture

$$\frac{d\overrightarrow{M}}{dt} = \frac{ds}{dt}\frac{d\overrightarrow{M}}{ds}$$

[9] Ce qui peut être ambigu si l'arc n'est pas simple.

pour représenter l'égalité s'écrivant plus rigoureusement $\overrightarrow{M}'(t) = s'(t)\ \overrightarrow{P}'(s(t))$ si la fonction P est définie sur J par $M(t) = P(s(t))$ pour tout $t \in I$.

L'intérêt de la notation différentielle est évident lorsqu'on fait de tels changements de variables directs et réciproques : si $I \ni t \mapsto M(t)$ est un paramétrage d'un arc régulier, tout paramétrage normal (noté abusivement $J \ni s \mapsto M(s)$) respectant l'orientation vérifiera

$$ds = \left\|\overrightarrow{M}'(t)\right\| dt$$

égalité dans laquelle peu nous importe que s soit fonction de t ou t fonction de s.

Si $M(t)$ est donné par ses coordonnées cartésiennes $(x(t), y(t))$ dans un repère orthonormal fixe, ceci pourra s'écrire

$$ds = \sqrt{x'^2(t) + y'^2(t)}\, dt$$

Si $M(t)$ est donné par un système de coordonnées polaires $(r(t), \theta(t))$ (avec r et θ de classe C^k sur I), nous aurons avec les notations usuelles

$$\overrightarrow{M}'(t) = r'(t)\ \overrightarrow{u}(\theta(t)) + r(t)\theta'(t)\overrightarrow{v}(\theta(t))$$

et donc

$$ds = \sqrt{r'^2(t) + r^2(t)\theta'^2(t)}\, dt$$

Lorsque l'arc est donné par une équation polaire $r = r(\theta)$, ceci deviendra évidemment

$$ds = \sqrt{r'^2(\theta) + r^2(\theta)}\, d\theta$$

Un paramétrage normal d'un arc régulier correspond à un parcours du support à une "vitesse numérique" constante (égale à 1). C'est un paramétrage "intrinsèque", qui va nous permettre de définir la notion géométrique de courbure d'un arc.

EXERCICE 20-2.15 Calculer la longueur du support de la cardioïde d'équation polaire

$$r = a(1 + \cos\theta)$$

20-2.2.2 Repère de Frenet

Nous considérons dans cette section un arc régulier Γ de classe C^k donné par le paramétrage $I \ni t \mapsto M(t)$ et nous considérons un paramétrage normal équivalent conservant l'orientation $J \ni s \mapsto M(s)$ avec

$$ds = \left\|\overrightarrow{M}'(t)\right\| dt$$

Le plan euclidien $\mathcal{E}_2$ est orienté.

DÉFINITION 20-2.16 *Soit $t_0 \in I$ et $s_0 \in J$ correspondant. On appelle vecteur unitaire de la tangente en t_0 à Γ (orientée dans le sens des t croissants) le vecteur*

$$\vec{T}(t_0) = \frac{\vec{M}'(t_0)}{\left\|\vec{M}'(t_0)\right\|}$$

Avec les notations introduites à la section précédente, nous avons

$$\vec{T}(t_0) = \frac{dt}{ds}(s_0)\,\frac{d\vec{M}}{dt}(t_0) = \frac{d\vec{M}}{ds}(s_0)$$

DÉFINITION 20-2.17 *Le vecteur normal principal $\vec{N}(t_0)$ est le vecteur unitaire se déduisant de $\vec{T}(t_0)$ par rotation d'angle $+\frac{\pi}{2}$. Le repère orthonormé direct*

$$\mathcal{R}_{t_0} = \left(M(t_0), \vec{T}(t_0), \vec{N}(t_0)\right)$$

est appelé repère de Frenet au point de paramètre t_0 de l'arc Γ.

Ce repère sera évidemment utilisé pour étudier localement l'arc, au voisinage de t_0. En travaillant dans un repère orthonormé direct $\left(O, \vec{\imath}, \vec{\jmath}\right)$ où les coordonnées de $M(t)$ sont $(x(t), y(t))$, nous aurons donc

$$\begin{cases} \vec{T}(t_0) = \dfrac{x'(t_0)}{\sqrt{x'^2(t_0) + y'^2(t_0)}}\vec{\imath} + \dfrac{y'(t_0)}{\sqrt{x'^2(t_0) + y'^2(t_0)}}\vec{\jmath} \\ \vec{N}(t_0) = -\dfrac{y'(t_0)}{\sqrt{x'^2(t_0) + y'^2(t_0)}}\vec{\imath} + \dfrac{x'(t_0)}{\sqrt{x'^2(t_0) + y'^2(t_0)}}\vec{\jmath} \end{cases}$$

On définit ainsi deux applications

$$I \ni t \mapsto \vec{T}(t) \quad \text{et} \quad I \ni t \mapsto \vec{N}(t)$$

qui sont évidemment de classe C^{k-1}. Avec toujours le même abus de notation, nous pourrons aussi considérer les applications (de classe C^{k-1} également)

$$J \ni s \mapsto \vec{T}(s) \quad \text{et} \quad J \ni s \mapsto \vec{N}(s)$$

avec cette fois

$$\vec{T}(s) = \frac{dx}{ds}\vec{\imath} + \frac{dy}{ds}\vec{\jmath} \quad \text{et} \quad \vec{N}(s) = -\frac{dy}{ds}\vec{\imath} + \frac{dx}{ds}\vec{\jmath}$$

et évidemment

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 1$$

sur J, puisque le paramétrage est normal.

20-2.2.3 Courbure, rayon de courbure en un point birégulier

Nous considérons dans cette section un arc régulier de classe C^k avec $k \geqslant 2$, paramétré par $I \ni t \mapsto M(t)$ et un paramétrage normal $J \ni s \mapsto M(s)$ correspondant à la même orientation. Nous étudions l'application de classe C^{k-1} définie par

$$J \ni s \mapsto \left(\overrightarrow{T}(s), \overrightarrow{N}(s)\right)$$

Nous allons étudier la dérivée de cette application ("repère mobile"), ce qui va nous amener aux **formules de Frenet**. Comme il s'agit de dériver une base orthonormale variable, nous savons que va apparaître un opérateur antisymétrique de l'espace vectoriel euclidien E_2 (*cf.* exercice **14-4**.9).

1ère formule de Frenet : pour tout $s \in J$, nous avons

$$\overrightarrow{T}^2(s) = 1$$

et on obtient en dérivant

$$\forall s \in J \quad \overrightarrow{T}(s) . \frac{d\overrightarrow{T}}{ds}(s) = 0$$

Le vecteur dérivé $\frac{d\overrightarrow{T}}{ds}(s)$ est donc colinéaire à $\overrightarrow{N}(s)$, deuxième vecteur du repère de Frenet en s. On définit donc la **courbure (algébrique)** $c(s)$ **de l'arc au point de paramètre** s par

$$\boxed{\frac{d\overrightarrow{T}}{ds}(s) = c(s)\,\overrightarrow{N}(s)}$$

2ème formule de Frenet : de même, le vecteur $\frac{d\overrightarrow{N}}{ds}(s)$ est colinéaire au vecteur $\overrightarrow{T}(s)$, et en dérivant l'égalité

$$\forall s \in J \quad \overrightarrow{T}(s) . \overrightarrow{N}(s) = 0$$

on obtient facilement

$$\boxed{\frac{d\overrightarrow{N}}{ds}(s) = -c(s)\,\overrightarrow{T}(s)}$$

L'interprétation géométrique de la courbure est simple si l'on fait intervenir une détermination de classe C^{k-1} de l'angle polaire de la tangente dans un repère fixe. Pour cela, nous utilisons le théorème de relèvement (*cf.* théorème **10-1**.56) appliqué à la fonction

$$z : J \to \mathbb{C} \quad s \mapsto z(s) = x_1(s) + i\,y_1(s)$$

où $x_1(s)$ et $y_1(s)$ sont les coordonnées de $\overrightarrow{T}(s)$ dans la base orthonormale directe fixe $\left(\overrightarrow{I}, \overrightarrow{J}\right)$. Avec l'abus d'écriture introduit plus haut, on a

$$z(s) = \frac{dx}{ds}(s) + i\,\frac{dy}{ds}(s)$$

L'application z est de classe $\mathcal{C}^{k-1}$, à valeurs dans l'ensemble des nombres complexes de module 1. Le théorème de relèvement nous assure l'existence d'une fonction $\alpha : J \to \mathbb{R}$ de classe $\mathcal{C}^{k-1}$ telle que

$$\forall s \in J \quad z(s) = e^{i\alpha(s)}$$

ce qui revient à écrire

$$\forall s \in J \quad \overrightarrow{T}(s) = \cos(\alpha(s))\,\overrightarrow{I} + \sin(\alpha(s))\,\overrightarrow{J}$$

soit encore

$$\alpha(s) = \left(\widehat{\overrightarrow{I}, \overrightarrow{T}(s)}\right) \quad \operatorname{mod} 2\pi$$

On a alors

$$\forall s \in J \quad \frac{d\overrightarrow{T}}{ds}(s) = \alpha'(s)\left(-\sin(\alpha(s))\,\overrightarrow{I} + \cos(\alpha(s))\,\overrightarrow{J}\right) = \alpha'(s)\,\overrightarrow{N}(s)$$

En comparant avec la première formule de Frenet, nous avons :

PROPOSITION 20-2.18 Si $J \ni s \mapsto M(s)$ est un paramétrage normal d'un arc régulier de classe $\mathcal{C}^k$ ($k \geqslant 2$), la courbure algébrique est l'application

$$c : J \ni s \mapsto c(s) = \frac{d\alpha}{ds}(s)$$

où $s \mapsto \alpha(s)$ est une détermination (de classe $\mathcal{C}^{k-1}$) de l'angle polaire du vecteur unitaire de la tangente orientée

$$\alpha(s) = \left(\widehat{\overrightarrow{I}, \overrightarrow{T}(s)}\right) \quad \operatorname{mod} 2\pi$$

Au plus cette courbure est grande en valeur absolue, au plus la variation de l'angle polaire de la tangente est grande, pour un "petit déplacement" donné sur l'arc.

La première formule de Frenet nous donne

$$\forall s \in J \quad \frac{d^2\overrightarrow{M}}{ds^2}(s) = \frac{d\overrightarrow{T}}{ds}(s) = c(s)\,\overrightarrow{N}(s)$$

et donc

$$c(s_0) = 0 \Leftrightarrow \left(\frac{d\overrightarrow{M}}{ds}(s_0), \frac{d^2\overrightarrow{M}}{ds^2}(s_0)\right) \text{ sont liés}$$
$$\Leftrightarrow s_0 \text{ n'est pas un point birégulier de l'arc}$$

EXEMPLE 20-2.19 Pour le cercle de centre O et de rayon R, considéré comme arc $\mathcal{C}^\infty$ dont un paramétrage admissible dans le repère $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$ est donné par

$$x(t) = R\cos t \quad \text{et} \quad y(t) = R\sin t$$

l'abscisse curviligne avec origine en $t = 0$ est évidemment

$$s(t) = Rt$$

et on peut prendre comme angle polaire de la tangente

$$\alpha(t) = t + \frac{\pi}{2}$$

Dans ce cas, la courbure est constante et vaut

$$\frac{d\alpha}{ds} = \frac{1}{R}$$

On est ainsi amené à définir **le rayon de courbure et le centre de courbure** en tout point **birégulier** d'un arc de classe C^k $(k \geqslant 2)$:

DÉFINITION 20-2.20 *En tout point birégulier d'un arc de classe C^k $(k \geqslant 2)$, le rayon de courbure algébrique est l'inverse de la courbure*

$$R(s) = \frac{1}{c(s)}$$

et le centre de courbure au point de paramètre s est

$$C(s) = M(s) + R(s)\ \overrightarrow{N}(s)$$

si $\left(\overrightarrow{T}(s), \overrightarrow{N}(s)\right)$ est le repère de Frenet en ce point. Le cercle de centre $C(s)$ et de rayon $|R(s)|$ est appelé cercle de courbure en s.

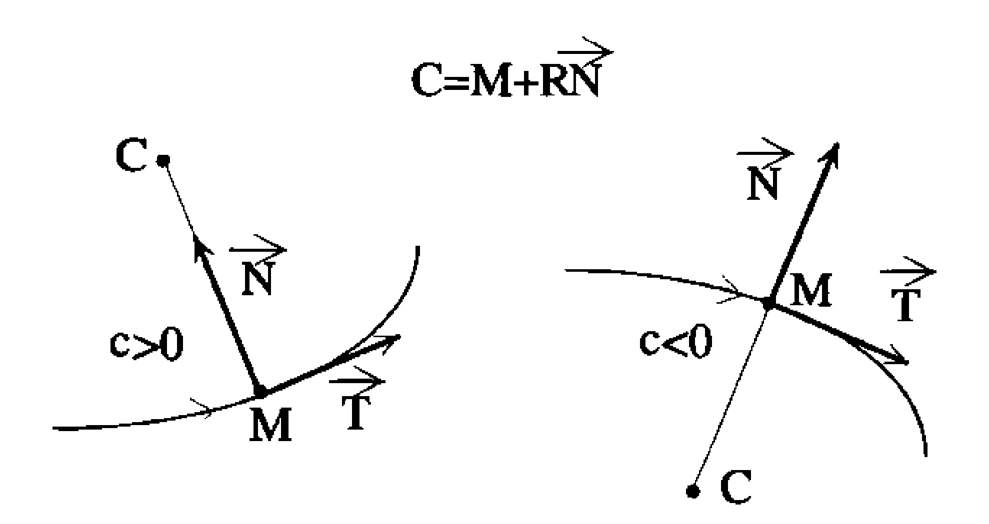


Figure **20.12** – *Repère de Frenet et signe de la courbure*

La courbure étudiée ici est algébrique. Nous avons interprété géométriquement sa valeur absolue et il reste à voir à quoi correspond son signe, qui dépend en fait de l'orientation choisie pour le plan :

Supposons donc l'arc étudié **birégulier** et le plan orienté comme à l'habitude, la rotation de $+\frac{\pi}{2}$ se faisant "dans le sens des aiguilles d'une montre". En tout point s_0 de J, nous savons que le vecteur $\frac{d^2\overrightarrow{M}}{ds^2}(s_0)$ est normal à l'arc et détermine sa concavité en s_0.

- Si $c(s_0) > 0$ alors $\frac{d^2\overrightarrow{M}}{ds^2}(s_0)$ et $\overrightarrow{N}(s_0)$ sont colinéaires et de même sens. La concavité de l'arc est donc située "à gauche" pour un observateur en $M(s_0)$ se déplaçant sur le support de l'arc dans le sens des s croissants.
- Par contre, si $c(s_0) < 0$, le vecteur $\overrightarrow{N}(s_0)$ n'est pas situé dans la concavité de l'arc, et celle-ci est donc "à droite" de l'observateur.
- Comme $R(s_0)$ et $c(s_0)$ sont de même signe, il en résulte que le centre de courbure $C(s_0)$ est toujours situé dans la concavité de l'arc en s_0.
- Un changement d'orientation dans le plan transforme le vecteur $\overrightarrow{N}(s_0)$ en son opposé, et change évidemment le signe de la courbure.
- L'utilisation d'un paramétrage normal changeant l'orientation de l'arc revient (à une translation près sur le paramètre) à paramétrer l'arc par $-s$. Il est clair que le repère de Frenet $\left(\overrightarrow{T}, \overrightarrow{N}\right)$ devient alors $\left(-\overrightarrow{T}, -\overrightarrow{N}\right)$, alors que $\frac{d\overrightarrow{T}}{ds} = \frac{d^2\overrightarrow{M}}{ds^2}$ ne change pas. Pour conserver la première formule de Frenet, il faut multiplier la courbure par -1 : **un changement d'orientation de l'arc change le signe de la courbure. Le centre de courbure ne change pas.**

DÉFINITION 20-2.21 *Lorsque le rayon de courbure algébrique est défini, sa valeur absolue est appelée rayon de courbure géométrique. Ce nombre ne dépend pas du paramétrage utilisé.*

20-2.3 Calcul pratique de la courbure

20-2.3.1 Formules générales

La formule de définition de la courbure fait intervenir un paramétrage normal. Dans la pratique, nous travaillons avec un paramétrage admissible de l'arc

$$I \ni t \mapsto M(t)$$

supposé régulier (ou birégulier si on veut calculer des rayons de courbure). Si $J \ni s \mapsto M(s)$ est un paramétrage normal conservant la même orientation, nous savons que (en notant comme d'habitude $[\overrightarrow{x}, \overrightarrow{y}]$ le produit mixte des vecteurs $\overrightarrow{x}$ et $\overrightarrow{y}$)

$$c(s_0) = \left[\overrightarrow{T}(s_0), c(s_0)\overrightarrow{N}(s_0)\right] = \left[\frac{d\overrightarrow{M}}{ds}(s_0), \frac{d^2\overrightarrow{M}}{ds^2}(s_0)\right]$$

On a au point $t_0 \in I$ correspondant, par dérivation de fonctions composées,

$$\begin{cases} \overrightarrow{M}'(t_0) = \dfrac{d\overrightarrow{M}}{dt}(t_0) = \dfrac{ds}{dt}(t_0)\,\dfrac{d\overrightarrow{M}}{ds}(s_0) = \left\|\overrightarrow{M}'(t_0)\right\|\,\dfrac{d\overrightarrow{M}}{ds}(s_0) \\ \overrightarrow{M}''(t_0) = \dfrac{d^2s}{dt^2}(t_0)\,.\dfrac{d\overrightarrow{M}}{ds}(s_0) + \left(\dfrac{ds}{dt}(t_0)\right)^2\dfrac{d^2\overrightarrow{M}}{ds^2}(s_0) \end{cases}$$

(nous donnerons à la section **20-2.6.2** une interprétation cinématique de ces égalités). On en déduit

$$\begin{aligned}&\left[\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\right] \\ &= \left[\frac{d\overrightarrow{M}}{ds}(s_0), \frac{d^2\overrightarrow{M}}{ds^2}(s_0)\right]\left(\frac{ds}{dt}\right)^3(t_0) = \left[\frac{d\overrightarrow{M}}{ds}(s_0), \frac{d^2\overrightarrow{M}}{ds^2}(s_0)\right]\left\|\overrightarrow{M}'(t_0)\right\|^3\end{aligned}$$

PROPOSITION 20-2.22 Si $\Gamma : I \ni t \mapsto M(t)$ est un arc régulier de classe C^k (avec $k \geqslant 2$), la courbure algébrique au point de paramètre t est donnée par

$$c(t) = \frac{\left[\overrightarrow{M}'(t), \overrightarrow{M}''(t)\right]}{\left\|\overrightarrow{M}'(t)\right\|^3}$$

En tout point birégulier, le rayon de courbure (algébrique) est donc donné par

$$R(t) = \frac{\left\|\overrightarrow{M}'(t)\right\|^3}{\left[\overrightarrow{M}'(t), \overrightarrow{M}''(t)\right]}$$

- **Calculs en coordonnées cartésiennes, dans un repère orthonormé $\left(O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right)$:**

  Si $(x(t), y(t))$ sont les coordonnées du point $M(t)$, nous avons

$$R(t) = \frac{\left(x'^2(t) + y'^2(t)\right)^{\frac{3}{2}}}{x'(t)\,y''(t) - y'(t)\,x''(t)}$$

  et comme le second vecteur du repère de Frenet au point de paramètre t est

$$\overrightarrow{N}(t) = -\frac{y'(t)}{\sqrt{x'^2(t) + y'^2(t)}}\overrightarrow{\imath} + \frac{x'(t)}{\sqrt{x'^2(t) + y'^2(t)}}\overrightarrow{\jmath}$$

  les coordonnées du centre de courbure $C(t) = M(t) + R(t)\,\overrightarrow{N}(t)$ seront

$$x_{C(t)} = x(t) - \frac{\left(x'^2(t) + y'^2(t)\right)y'(t)}{x'(t)\,y''(t) - y'(t)\,x''(t)}$$

$$\text{et} \quad y_{C(t)} = y(t) + \frac{\left(x'^2(t) + y'^2(t)\right)x'(t)}{x'(t)\,y''(t) - y'(t)\,x''(t)}$$

  (ces formules sont compliquées, on remarquera cependant la disparition des radicaux).

- **Dans le cas particulier d'un arc admettant une équation cartésienne** $y = f(x)$ avec f de classe $\mathcal{C}^k$:

  L'abscisse x est un paramètre admissible, et la formule précédente donne
  $$R(x) = \frac{(1 + f'^2(x))^{\frac{3}{2}}}{f''(x)}$$

- **Pour un arc donné par une équation polaire** $r = r(\theta)$:

  Dans ce cas $\overrightarrow{OM}(\theta) = r(\theta)\overrightarrow{u}(\theta)$ donne $\overrightarrow{M}'(\theta) = r'(\theta)\overrightarrow{u}(\theta) + r(\theta)\overrightarrow{v}(\theta)$ et donc
  $$\overrightarrow{T}(\theta) = \frac{r'(\theta)\overrightarrow{u}(\theta) + r(\theta)\overrightarrow{v}(\theta)}{\sqrt{r^2(\theta) + r'^2(\theta)}} \text{ et } \overrightarrow{N}(\theta) = \frac{-r(\theta)\overrightarrow{u}(\theta) + r'(\theta)\overrightarrow{v}(\theta)}{\sqrt{r^2(\theta) + r'^2(\theta)}}$$
  avec
  $$R(\theta) = \frac{(r^2(\theta) + r'^2(\theta))^{\frac{3}{2}}}{r^2(\theta) + 2r'^2(\theta) - r(\theta)r''(\theta)}$$
  Lorsque l'arc ne passe pas par l'origine et si l'expression
  $$q(\theta) = \frac{1}{r(\theta)}$$
  est plus simple à dériver, on pourra remarquer que $R(\theta)$ s'écrit aussi
  $$R(\theta) = \frac{(q^2(\theta) + q'^2(\theta))^{\frac{3}{2}}}{(q^3(\theta))(q(\theta) + q''(\theta))}$$

EXERCICE 20-2.23 Soit F un point du plan euclidien et Δ une droite du plan passant par F. On considère un arc birégulier Γ possédant la propriété suivante : la normale en un point M quelconque coupe Δ en P, la perpendiculaire à MP en P coupe FM en Q et la perpendiculaire à FM en Q coupe la normale en un point Ω qui est le centre de courbure en M. Montrer que le support de Γ est inclus dans une conique de foyer F et d'axe focal Δ.

20-2.3.2 Méthodes pratiques

On préfère souvent appliquer la définition
$$c(s) = \frac{d\alpha}{ds}(s)$$
plutôt que les formules précédentes, lorsque
$$ds = \left\|\overrightarrow{M}'(t)\right\| dt$$
est "simple" et lorsqu'on accède facilement à la dérivée de la fonction $t \mapsto \alpha(t)$, soit parce que $\alpha(t)$ est une expression simple de t ou parce qu'on dérive simplement une des expressions
$$\cos\alpha(t) = \frac{x'(t)}{\sqrt{x'^2(t) + y'^2(t)}}, \quad \sin\alpha(t) = \frac{y'(t)}{\sqrt{x'^2(t) + y'^2(t)}} \text{ ou } \tan\alpha(t) = \frac{y'(t)}{x'(t)}$$

EXERCICE 20-2.24 On considère l'arc paramétré défini dans un repère orthonormé $\left(O, \vec{\imath}, \vec{\jmath}\right)$ par

$$x(t) = a(t - \sin t) \quad \text{et} \quad y(t) = a(1 - \cos t)$$

avec $t \in]0, 2\pi[$ ($a > 0$ est fixé).

Montrer que cet arc est régulier, et que son support est trajectoire d'un point fixe d'un cercle de rayon a qui roule sans glisser sur l'axe Ox (cycloïde droite). Montrer que le lieu du centre de courbure se déduit de ce support par des transformations géométriques simples.

Résultats : grâce aux formules de trigonométrie, on a

$$\begin{cases} x'(t) = 2a \sin^2 \dfrac{t}{2} \\ y'(t) = 2a \sin \dfrac{t}{2} \cos \dfrac{t}{2} \end{cases}$$

On en déduit immédiatement, (pour $t \in]0, 2\pi[$, $\sin \frac{t}{2} > 0$)

$$ds = 2a \sin \frac{t}{2}\, dt \quad \text{et} \quad \vec{T}(t) = \sin \frac{t}{2} \vec{\imath} + \cos \frac{t}{2} \vec{\jmath}$$

On peut donc choisir

$$\alpha(t) = \frac{\pi}{2} - \frac{t}{2}$$

et donc

$$R(t) = -4a \sin \frac{t}{2}$$

Le centre de courbure en $M(t)$ a pour coordonnées

$$x_{C(t)} = a(t + \sin t) \quad \text{et} \quad y_{C(t)} = -a(1 - \cos t)$$

Pour un arc défini par une équation polaire $r = r(\theta)$, on détermine en général l'angle orienté des droites $OM(\theta)$ et T_θ tangente au point de paramètre θ par

$$\tan V(\theta) = \frac{r(\theta)}{r'(\theta)}$$

Comme

$$\alpha(\theta) = V(\theta) + \theta \mod \pi$$

la fonction $\theta \mapsto \alpha(\theta) - (V(\theta) + \theta)$ est constante sur tout intervalle, ce qui donnera, avec la notation différentielle

$$d\alpha = dV + d\theta$$

EXERCICE 20-2.25 On appelle spirale logarithmique tout arc admettant, dans un repère orthonormé bien choisi, une équation polaire de la forme

$$r(\theta) = a\, e^{m\theta}$$

avec $a > 0$ et $m \neq 0$. Montrer que le lieu du centre de courbure à cet arc est également une spirale logarithmique.

20-2.3.3 Courbure à l'origine

Un repère orthonormé direct $\left(O, \overrightarrow{I}, \overrightarrow{J}\right)$ étant donné, considérons un arc paramétré de classe C^k régulier $t \mapsto M(t)$ défini sur un voisinage de 0, vérifiant

$$M(0) = O \quad \text{et} \quad \overrightarrow{M}'(0) \text{ colinéaire à } \overrightarrow{I}$$

La tangente à l'origine est donc l'axe des abscisses. Comme on a $x(0) = 0$ et $x'(0) \neq 0$, l'application $t \mapsto x(t)$ définit un C^k-difféomorphisme entre deux intervalles ouverts contenant chacun 0. L'abscisse x est donc un paramètre admissible au voisinage de 0, et l'arc considéré admet une équation cartésienne de la forme

$$y = f(x)$$

avec f de classe C^k au voisinage de 0. On a $f(0) = f'(0) = 0$, puisque la tangente à l'origine est horizontale. D'après les résultats de la section **20-2.3.1**, la courbure au point d'abscisse x vaut

$$c(x) = \frac{f''(x)}{\left(1 + f'^2(x)\right)^{\frac{3}{2}}}$$

et en particulier

$$c(0) = f''(0)$$

Si le point d'abscisse 0 est birégulier, le rayon de courbure à l'origine sera donné par

$$R_O = \frac{1}{f''(0)}$$

Dans la pratique, on ne connaît pas explicitement la fonction f. Mais la formule de Taylor-Young donne, au voisinage de 0,

$$y = f(x) = \frac{1}{2} f''(0)\, x^2 + o\left(x^2\right)$$

et par conséquent

$$R_O = \frac{1}{f''(0)} = \lim_{t \to 0} \frac{x^2(t)}{2y(t)}$$

On utilisera cette technique pour déterminer la courbure en un point particulier M_0 d'un arc paramétré ou d'une courbe définie par une équation cartésienne. On pourra faire un changement de repère pour amener M_0 à l'origine et la tangente en ce point sur l'axe des abscisses, et déterminer dans ce repère (soit en utilisant un paramétrage, soit en utilisant une équation cartésienne) la limite du quotient $\dfrac{X^2}{2Y}$ lorsqu'un point du support $M(X, Y)$ tend vers O :

EXERCICE 20-2.26 On considère l'ellipse d'équation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

dans un repère orthonormé. Montrer que le rayon de courbure géométrique au sommet $A(a,0)$ vaut

$$R_A=\frac{b^2}{a}$$

EXERCICE 20-2.27 On considère un arc d'équation polaire

$$r=f(\theta)$$

avec f de classe C^k au voisinage de θ_0. On suppose que $f(\theta_0)=0$ et $f'(\theta_0)\neq 0$. On oriente l'arc de manière à ce que le repère de Frenet en θ_0 soit

$$\mathcal{R}_{\theta_0}=(O,\overrightarrow{u}(\theta_0),\overrightarrow{v}(\theta_0))$$

Montrer que le rayon de courbure en θ_0 vaut

$$R_O=\frac{f'(\theta_0)}{2}$$

20-2.4 Cercle de courbure

En un point birégulier t_0 d'un arc de classe C^k, nous avons défini le centre de courbure

$$C(t_0)=M(t_0)+R(t_0)\,\overrightarrow{N}(t_0)$$

et le cercle de courbure, de centre $C(t_0)$ et de rayon $|R(t_0)|$. Ce dernier apparaît de manière naturelle lorsqu'on fait une étude locale de position relative de $M(t)$ par rapport à un cercle arbitraire passant par $M(t_0)$. Il s'agit d'une étude analogue à celle de la section **20-1.2.4** (où l'on travaillait avec une droite variable passant par $M(t_0)$) :

Choisissons un repère orthonormé direct $\left(O,\overrightarrow{I},\overrightarrow{J}\right)$ avec $O=M(t_0)$ de manière à ce que la tangente à l'arc en t_0 soit dirigée par $\overrightarrow{I}$. Comme nous l'avons vu dans la section précédente, on peut localement paramétrer l'arc par son abscisse x pour lui donner une représentation cartésienne

$$y=f(x)$$

avec f de classe C^k au voisinage de 0, vérifiant $f(0)=f'(0)=0$ et $f''(0)\neq 0$, puisque l'arc est supposé birégulier. Considérons un cercle C passant par O et de centre $\Omega(\alpha,\beta)$. Son équation dans $\left(O,\overrightarrow{I},\overrightarrow{J}\right)$ est

$$Q(x,y)=x^2+y^2-2\alpha x-2\beta y=0$$

Nous savons que si P est un point quelconque du plan, de coordonnées (x,y), la quantité $Q(x,y)$ représente

$$Q(x,y)=\left\|\overrightarrow{\Omega P}\right\|^2-\left(\alpha^2+\beta^2\right)=\left\|\overrightarrow{\Omega P}\right\|^2-R^2$$

où R est le rayon du cercle et donc le signe de $Q(x,y)$ change selon que P est intérieur ou extérieur au cercle $\mathcal{C}$.

Pour placer, pour x proche de 0, le point $M(x, f(x))$ par rapport au cercle $\mathcal{C}$, on étudie le signe de

$$\Delta(x) = x^2 + f^2(x) - 2\alpha x - 2\beta f(x)$$

(en prenant un équivalent). Pour x tendant vers 0, $f(x)$ est infiniment petit d'ordre 2. En supposant par exemple f de classe $\mathcal{C}^3$, la formule de Taylor-Young appliquée à f donne

$$\Delta(x) = -2\alpha x + (1 - \beta f''(0)) x^2 - \beta \frac{f'''(0)}{3} x^3 + o(x^3)$$

La discussion se mène en fonction de la position de Ω :

- Si $\alpha \neq 0$, c'est à dire si Ω **n'est pas sur la normale à l'arc en** O :
  Le cercle $\mathcal{C}$ est alors "sécant" et
  $$\Delta(x) \sim -2\alpha x$$
  change de signe avec x : **l'arc traverse le cercle en** O.

- Si $\alpha = 0$ et $\beta \neq \dfrac{1}{f''(0)} = R_O$, Ω n'est pas le centre de courbure en O : **le cercle** $\mathcal{C}$ **est tangent à l'arc, mais n'est pas le cercle de courbure :**
  $$\Delta(x) \sim (1 - \beta f''(0)) x^2$$
  garde un signe constant : **l'arc reste localement du même côté du cercle.**

- Si $\alpha = 0$ et $\beta = \dfrac{1}{f''(0)} = R_O$, Ω est le centre de courbure en O : **le cercle** $\mathcal{C}$ **est le cercle de courbure :**
  $$\Delta(x) = -\frac{f'''(0)}{3f''(0)} x^3 + o(x^3)$$
  Lorsque $f'''(0) \neq 0$, $\Delta(x)$ est donc un infiniment petit d'ordre 3 et **l'arc traverse son cercle de courbure en** O. On dit que le cercle de courbure est **osculateur** (surosculateur si $\Delta(x) = o(x^3)$) (figure **20.13**).

REMARQUE 20-2.28 On arriverait à la même conclusion en faisant un calcul intrinsèque : si $s \mapsto M(s)$ est un paramétrage normal (défini au voisinage de 0) d'un arc birégulier de classe $\mathcal{C}^k$ (avec $k \geqslant 3$, ce qui assure le caractère $\mathcal{C}^1$ de l'application $s \mapsto R(s)$), les formules de Frenet et la formule de Taylor-Young donnent, pour s tendant vers 0

$$\overrightarrow{M(0)\,M(s)} = s\,\overrightarrow{T}_0 + \frac{s^2}{2}\frac{\overrightarrow{N}_0}{R_0} + \frac{s^3}{6}\left(-\frac{\overrightarrow{T}_0}{R_0^2} - \frac{R_0'}{R_0^2}\overrightarrow{N}_0\right) + o(s^3)$$

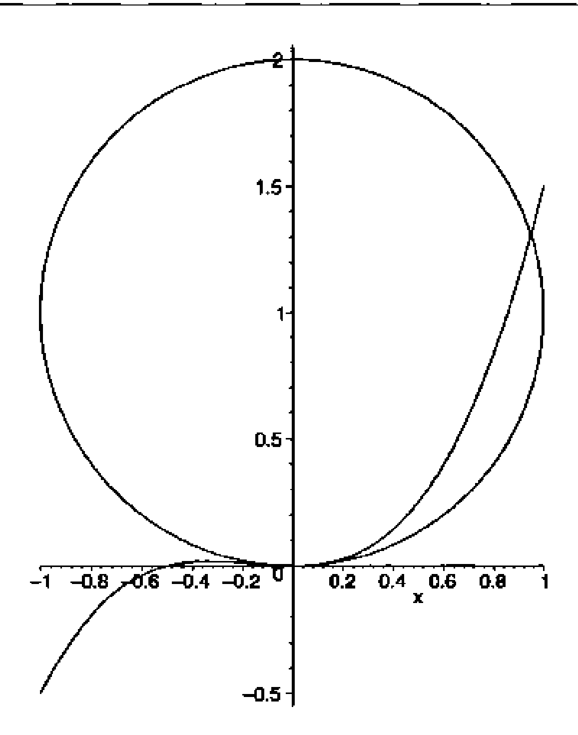


Figure **20.13** – *Positions relatives de la courbe* $y = \dfrac{x^2}{2} + x^3$ *et de son cercle de courbure à l'origine*

où $\left(\overrightarrow{T}_0, \overrightarrow{N}_0\right)$ est le repère de Frenet en 0 et (R_0, R_0') représentent le rayon de courbure ainsi que sa dérivée en 0.

La position de $M(s)$ par rapport au cercle de courbure en 0 est donnée par le signe de $\delta(s) = \left\|\overrightarrow{C(0)\,M(s)}\right\|^2 - R_0^2$, soit

$$\delta(s) = \left\| -R_0 \overrightarrow{N}_0 + s\,\overrightarrow{T}_0 + \frac{s^2}{2}\,\frac{\overrightarrow{N}_0}{R_0} + \frac{s^3}{6}\left(-\frac{\overrightarrow{T}_0}{R_0^2} - \frac{R_0'}{R_0^2}\overrightarrow{N}_0\right) + o\left(s^3\right)\right\|^2 - R_0^2$$

Le développement limité à l'ordre 3 de cette quantité s'écrit

$$\delta(s) = \frac{R_0'}{6R_0} s^3 + o\left(s^3\right)$$

Nous retrouvons en général un infiniment petit d'ordre 3. On peut également retenir que, chaque fois que le rayon de courbure passe par un extremum (R' s'annule), le cercle de courbure est surosculateur.

20-2.5 Développée d'un arc plan

Considérons un arc birégulier de classe C^k (avec $k \geqslant 3$)

$$\Gamma : I \ni t \mapsto M(t)$$

On appelle développée de Γ l'arc de classe C^{k-2}

$$\gamma : I \ni t \mapsto C(t) = M(t) + R(t)\,\overrightarrow{N}(t)$$

Son support est donc le lieu du centre de courbure en $M(t)$.

Si $J \ni s \mapsto M(s)$ est un paramétrage normal de Γ, on obtient un paramétrage admissible de γ

$$J \ni s \mapsto C(s) = M(s) + R(s)\ \overrightarrow{N}(s)$$

Les formules de Frenet donnent alors

$$\overrightarrow{C}'(s) = R'(s)\ \overrightarrow{N}(s)$$

Sur tout sous-intervalle de J (ou de I) sur lequel la dérivée du rayon de courbure ne s'annule pas, la développée est régulière. La tangente à la développée en s est

$$C(s) + \mathbb{R}\,\overrightarrow{N}(s)$$

c'est donc la normale à l'arc Γ en $M(s)$. Ceci nous permettrait de donner une deuxième définition de la développée de Γ comme enveloppe des normales à Γ.

REMARQUE 20-2.29 Sur un tel sous-intervalle J_1 de J, un vecteur unitaire de la tangente à la développée au point de paramètre s est $\overrightarrow{N}(s)$. On peut choisir un paramétrage normal de la développée $K_1 \ni \sigma \mapsto C(\sigma)$ de manière à ce que le repère de Frenet en tout point s_0 (correspondant à σ_0) soit égal à

$$\left(C(s_0), \overrightarrow{N}(s_0), -\overrightarrow{T}(s_0)\right)$$

On a alors

$$\forall s \in J_1 \quad R'(s)\ \overrightarrow{N}(s) = \frac{d\overrightarrow{C}}{ds}(s) = \frac{d\sigma}{ds}(s)\frac{d\overrightarrow{C}}{d\sigma}(\sigma) = \frac{d\sigma}{ds}(s)\ \overrightarrow{N}(s)$$

ce qui montre que

$$\forall s \in J_1 \quad \sigma(s) - R(s) \text{ est constant}$$

Cette remarque est utile pour rectifier rapidement la développée : pour s_1 et $s_2 \in J_1$, la longueur de la développée $\gamma|_{[s_1,s_2]}$ vaut simplement $|\sigma(s_2) - \sigma(s_1)|$, soit $|R(s_2) - R(s_1)|$.

On retiendra que la développée présente un point stationnaire lorsque la dérivée du rayon de courbure à l'arc s'annule (donc en particulier lorsque ce rayon de courbure passe par un extremum).

EXERCICE 20-2.30 Déterminer la développée de l'ellipse d'équation

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$$

Comment son support se déduit-il de celui de l'arc étudié à l'exercice 20-1.21 ? A quelle condition sur α et β la longueur de la développée vaut-elle 7α ?

EXERCICE 20-2.31 Montrer que la développée de la cardioïde d'équation polaire

$$r(\theta) = a(1 + \cos\theta)$$

est encore une cardioïde. Indication : il faudra choisir une nouvelle origine correspondant au point stationnaire sur la développée. Où est situé ce point ?

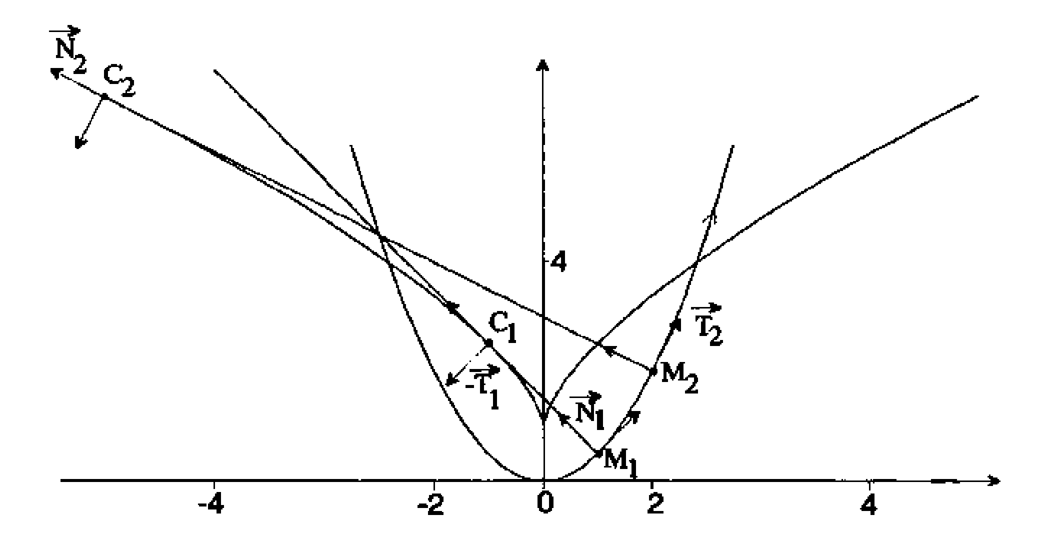


Figure **20.14** – *Développée de la parabole* $y = \dfrac{x^2}{2}$

EXERCICE 20-2.32 Si $\Gamma : I \ni t \mapsto M(t)$ est un arc birégulier de classe C^k (avec $k \geqslant 4$), tout arc paramétré de la forme

$$\Gamma_\alpha : I \ni t \mapsto P(t) = M(t) + \alpha \overrightarrow{N}(t)$$

où α est un réel fixé est appelé arc parallèle à Γ. Montrer que Γ_α est de classe C^{k-1} et qu'en tout point birégulier t (pour Γ_α), Γ et Γ_α ont même centre de courbure. Un arc et un arc parallèle ont donc en général même développée.

20-2.6 Courbure en dimension 3

20-2.6.1 Définition

Considérons **un arc régulier** de classe C^k ($k \geqslant 1$)

$$I \ni t \mapsto M(t) \in \mathcal{E}_3$$

dont le support est tracé dans un espace affine euclidien de dimension 3. La notion de paramétrage normal est définie comme dans le plan, les résultats de la section **20-2.2.1** n'étant pas spécifiques à la dimension 2. Un paramétrage normal $J \ni s \mapsto M(s)$ respectant l'orientation vérifie toujours

$$ds = \left\| \overrightarrow{M}'(t) \right\| dt$$

- **En coordonnées cartésiennes**, dans un repère orthonormé fixe,

$$ds = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}\, dt$$

- **En coordonnées cylindriques d'axe** Oz, on a

$$\overrightarrow{OM}(t) = r(t)\ \overrightarrow{u}(\theta(t)) + z(t)\ \overrightarrow{k}$$

et donc

$$\overrightarrow{M}'(t) = r'(t)\ \overrightarrow{u}(\theta(t)) + r(t)\ \theta'(t)\ \overrightarrow{v}(\theta(t)) + z'(t)\ \overrightarrow{k}$$

Comme la base $\left(\overrightarrow{u}(\theta), \overrightarrow{v}(\theta), \overrightarrow{k}\right)$ est orthonormale, on obtient

$$ds = \sqrt{r'^2(t) + r^2(t)\ \theta'^2(t) + z'^2(t)}\, dt$$

- **En coordonnées sphériques**, avec

$$\overrightarrow{OM}(t) = r(t)\ \cos\varphi(t)\ \overrightarrow{u}(\theta(t)) + r(t)\ \sin\varphi(t)\ \overrightarrow{k}$$

nous aurons

$$\begin{aligned}\overrightarrow{M}'(t) = r'(t)\ &\left(\cos\varphi(t)\ \overrightarrow{u}(\theta(t)) + \sin\varphi(t)\ \overrightarrow{k}\right)\\ &+ r(t)\ \varphi'(t)\ \left(-\sin\varphi(t)\ \overrightarrow{u}(\theta(t)) + \cos\varphi(t)\ \overrightarrow{k}\right)\\ &\qquad + r(t)\ \cos(\varphi(t))\ \theta'(t)\ \overrightarrow{v}(\theta(t))\end{aligned}$$

Le vecteur dérivé est ainsi décomposé dans une base orthonormale et donc

$$ds = \sqrt{r'^2(t) + r^2(t)\ \varphi'^2(t) + r^2(t)\ \theta'^2(t)\ \cos^2(\varphi(t))}\, dt$$

- Le vecteur unitaire de la tangente orientée au point de paramètre t, correspondant au paramètre normal s, est toujours

$$\overrightarrow{T}(t) = \frac{\overrightarrow{M}'(t)}{\left\|\overrightarrow{M}'(t)\right\|} = \frac{d\overrightarrow{M}}{ds}(s)$$

Pour un arc de classe C^k ($k \geqslant 2$), le vecteur

$$\frac{d^2\overrightarrow{M}}{ds^2}(s) = \frac{d\overrightarrow{T}}{ds}(s)$$

est un vecteur normal à $\overrightarrow{T}(s)$ (puisque la norme de $\overrightarrow{T}$ est constante) ; il est non nul si et seulement si le point de paramètre s est birégulier. Comme, sur une normale à $\overrightarrow{T}$, il n'y a pas de sens privilégié (pas d'orientation intrinsèque d'un plan en dimension 3), nous définirons la **courbure géométrique** en s par

$$c(s) = \left\|\frac{d^2\overrightarrow{M}}{ds^2}(s)\right\| = \left\|\frac{d\overrightarrow{T}}{ds}(s)\right\|$$

En un **point birégulier**, $c(s) \neq 0$, et le rayon de courbure géométrique est alors défini par

$$R(s) = \frac{1}{c(s)}$$

par analogie avec ce qui a été fait dans le plan.

En un tel point, le plan osculateur à l'arc au point s (correspondant au paramètre t pour l'arc de départ) est par définition

$$\begin{aligned}\mathcal{P}_t &= M(t) + \operatorname{vect}\left(\overrightarrow{M}'(t), \overrightarrow{M}''(t)\right) \\ &= M(s) + \operatorname{vect}\left(\overrightarrow{T}(s) = \frac{d\overrightarrow{M}}{ds}(s), \frac{d^2\overrightarrow{M}}{ds^2}(s) = \frac{d\overrightarrow{T}}{ds}(s)\right)\end{aligned}$$

Le vecteur non nul

$$\frac{d^2\overrightarrow{M}}{ds^2}(s) = \frac{d\overrightarrow{T}}{ds}(s)$$

définit le demi-plan osculateur (voir section **20-1.2.3**) et on définit le **vecteur normal principal** $\overrightarrow{N}(s)$ par

$$\frac{d\overrightarrow{T}}{ds}(s) = c(s)\,\overrightarrow{N}(s)$$

C'est donc un vecteur unitaire, normal à l'arc et définissant la concavité de celui-ci.

20-2.6.2 Calcul de la courbure, interprétation cinématique

On suppose l'espace orienté, ce qui permet d'y définir un produit vectoriel. Le calcul de la courbure se fait comme dans le plan, en remplaçant un produit mixte par la norme d'un produit vectoriel : en un point $t_0 \in I$ correspondant au paramètre normal s_0

$$\begin{cases}\overrightarrow{M}'(t_0) = \dfrac{d\overrightarrow{M}}{dt}(t_0) = \dfrac{ds}{dt}(t_0)\,\dfrac{d\overrightarrow{M}}{ds}(s_0) = \left\|\overrightarrow{M}'(t_0)\right\|\,\dfrac{d\overrightarrow{M}}{ds}(s_0) \\ \overrightarrow{M}''(t_0) = \dfrac{d^2s}{dt^2}(t_0).\dfrac{d\overrightarrow{M}}{ds}(s_0) + \left(\dfrac{ds}{dt}(t_0)\right)^2\dfrac{d^2\overrightarrow{M}}{ds^2}(s_0)\end{cases}$$

et donc

$$\overrightarrow{M}'(t_0) \wedge \overrightarrow{M}''(t_0) = \left(\frac{ds}{dt}(t_0)\right)^3\frac{d\overrightarrow{M}}{ds}(s_0) \wedge \frac{d^2\overrightarrow{M}}{ds^2}(s_0)$$

Comme les vecteurs $\dfrac{d\overrightarrow{M}}{ds}(s_0)$ et $\dfrac{d^2\overrightarrow{M}}{ds^2}(s_0)$ sont orthogonaux, nous avons

$$\left\|\overrightarrow{M}'(t_0) \wedge \overrightarrow{M}''(t_0)\right\| = \left|\left(\frac{ds}{dt}(t_0)\right)^3\right|\left\|\frac{d\overrightarrow{M}}{ds}(s_0)\right\|\left\|\frac{d^2\overrightarrow{M}}{ds^2}(s_0)\right\|$$

soit

$$\left\|\overrightarrow{M}'(t_0) \wedge \overrightarrow{M}''(t_0)\right\| = \left\|\overrightarrow{M}'(t_0)\right\|^3 c(s_0)$$

On en déduit qu'en tout point $t \in I$, la courbure est donnée par

$$c(t) = \frac{\left\| \overrightarrow{M}'(t) \wedge \overrightarrow{M}''(t) \right\|}{\left\| \overrightarrow{M}'(t) \right\|^3}$$

et en un point birégulier, le rayon de courbure vaut

$$R(t) = \frac{\left\| \overrightarrow{M}'(t) \right\|^3}{\left\| \overrightarrow{M}'(t) \wedge \overrightarrow{M}''(t) \right\|}$$

Si l'application $t \mapsto M(t)$ modélise un mouvement ponctuel, le vecteur $\dfrac{d\overrightarrow{M}}{dt}(t)$ représente le vecteur vitesse intantanée à l'instant t, et sa norme

$$v = \left\| \frac{d\overrightarrow{M}}{dt} \right\| = \frac{ds}{dt}$$

est la vitesse numérique. Le vecteur accélération

$$\frac{d^2\overrightarrow{M}}{dt^2} = \frac{d^2s}{dt^2}.\overrightarrow{T} + \left(\frac{ds}{dt}\right)^2 c\,\overrightarrow{N}$$

admet une composante tangentielle

$$\frac{d^2s}{dt^2}.\overrightarrow{T} = \frac{dv}{dt}\overrightarrow{T}$$

($\dfrac{dv}{dt}$ est appelée accélération numérique) et une composante normale

$$\left(\frac{ds}{dt}\right)^2 c\,\overrightarrow{N} = \frac{v^2}{R}\overrightarrow{N}$$

en tout point birégulier où le rayon de courbure est défini (en un point où c est nul, on pose $R = +\infty$, et l'accélération n'a pas de composante normale).

20-3 Exercices

EXERCICE 20-3.1 Dans le plan rapporté à un repère orthonormé on donne un triangle par les équations de ses côtés : $(\mathcal{D}_i)$ $a_i x + b_i y + c_i = 0$, $i = 1\ldots3$. Déterminer l'aire de ce triangle.

EXERCICE 20-3.2 Pour $t \in \mathbb{R}$, le plan étant rapporté à un repère orthonormé, on considère la droite d'équation $x + ty + at^3 = 0$, a est une constante réelle. Déterminer l'ensemble des points du plan où passent trois de ces droites, l'une étant bissectrice des deux autres.

EXERCICE 20-3.3 Déterminer l'image du cercle unité par la transformation

$$z \mapsto \frac{1}{(1-z)^2}$$

EXERCICE 20-3.4 Nature et définition géométrique de l'arc défini en repère orthonormé par

$$x(t) = \frac{1}{1+t+t^2} \; y(t) = \frac{t}{1+t+t^2}$$

EXERCICE 20-3.5 Déterminer le lieu des points d'où l'on voit la cardioïde d'équation polaire

$$r = a(1+\cos\theta)$$

sous un angle droit.

EXERCICE 20-3.6 Dans un repère orthonormé, Γ est la courbe d'équation polaire $r(\theta) = \dfrac{1}{\theta\cos\theta - \sin\theta}$. Montrer que Γ possède une infinité d'asymptotes qui sont toutes tangentes à une courbe algébrique simple. (On rappelle qu'une courbe algébrique est donnée dans un repère du plan par une équation cartésienne polynomiale).

EXERCICE 20-3.7 Déterminer un arc plan passant par O, dont la tangente en M coupe Ox en T tel que la longueur de l'arc OM soit égale à $2\|\overrightarrow{MT}\|$.

EXERCICE 20-3.8 Déterminer le rayon de courbure en O de la courbe définie par l'équation cartésienne dans un repère orthonormé : $x^4 + y^4 + x^3 + y^3 = 0$.

EXERCICE 20-3.9 Une conique est tangente à Ox en $A(a,0)$ et à Oy en $B(0,b)$. Montrer que les rayons de courbure en A et B vérifient $b^3R_A = a^3R_B$.

EXERCICE 20-3.10 Une cardioïde roule sans glisser sur une droite. Déterminer le lieu du point de rebroussement et du centre de courbure au point de contact avec la droite.

EXERCICE 20-3.11 Dans le plan affine euclidien muni d'un repère orthonormé $(O,\vec{\imath},\vec{\jmath})$, on considère le point $A(a,0)$, où a est un réel strictement positif.

1. P étant un point variable de l'axe $y'0y$, montrer que l'ensemble $\mathcal{F}$ des points M du plan tels que $\| \overrightarrow{PM} \| = a$ et $\overrightarrow{MP}.\overrightarrow{MA} = 0$ se décompose en une droite $(\mathcal{D})$ et une courbe (C) que l'on représentera.

2. Montrer que (C) est le support d'un arc paramétré γ de classe C^∞ et définir les centres de courbure à γ en 0 et A.

EXERCICE 20-3.12 Déterminer la perpendiculaire commune à la droite D d'équations, en repère orthonormé : $(x = a, y = b)$ et à la droite D' d'équations

$$(x + cy + z = 0, cx - y + z = 0)$$

EXERCICE 20-3.13 Déterminer l'aire de la portion de plan comprise entre la parabole d'équation

$$y^2 = 2px$$

et sa développée.

EXERCICE 20-3.14 On considère la courbe définie en coordonnées polaires par

$$\rho = \frac{a}{1+\cos\theta}$$

Quelle relation y-a-t-il entre V (angle entre l'horizontale et la tangente orientée) et θ ? Déterminer le lieu du projeté $Q(\theta)$ du pôle O sur la tangente au point de paramètre θ. A quelle condition les tangentes en θ_1 et θ_2 sont-elles orthogonales ? Déterminer alors le lieu de leur point d'intersection.

EXERCICE 20-3.15 On se place dans le plan euclidien et on considère F et F' deux points distincts. Soit $a > 0$.

1. Quel est l'ensemble $\mathcal{E}$ des points M vérifiant $MF + MF' = 2a$?
2. Montrer que la tangente en M à $\mathcal{E}$ est bissectrice extérieure à $\widehat{FMF'}$.
3. Déterminer le lieu des symétriques de F par rapport aux tangentes à $\mathcal{E}$?
4. Déterminer le lieu des projections orthogonales de F sur les tangentes ?

EXERCICE 20-3.16 Tracer la courbe d'équation $x^3+y^3-3xy = 0$ dans le plan rapporté à un repère orthonormé.

EXERCICE 20-3.17 Construire la courbe d'équation polaire

$$r = \frac{\sin 2\theta}{2\cos\theta\ - 1}$$

Chapitre 21

Surfaces

21-1 Nappes paramétrées

Dans tout ce chapitre $\mathcal{E}_3$ est un espace affine de dimension 3. Lorsque la notion d'orthogonalité est utilisée, on suppose de plus que $\mathcal{E}_3$ est euclidien orienté (ce qui permet d'utiliser le produit vectoriel). $\mathcal{R} = \left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$ est un repère affine fixe de $\mathcal{E}_3$.

21-1.1 Définition

DÉFINITION 21-1.1 *Soit $k \in \mathbb{N}^*$. Une nappe paramétrée de classe C^k est une application*

$$M : \mathcal{U} \to \mathcal{E}_3 \quad (u,v) \mapsto M(u,v)$$

de classe C^k définie sur un ouvert connexe $\mathcal{U}$ de $\mathbf{R}^2$.

Si on travaille dans le repère $\mathcal{R}$, la donnée de l'application M équivaut à celle de trois fonctions numériques

$$\mathcal{U} \ni (u,v) \mapsto x(u,v), \quad \mathcal{U} \ni (u,v) \mapsto y(u,v) \quad \text{et} \quad \mathcal{U} \ni (u,v) \mapsto z(u,v)$$

telles que

$$\forall (u,v) \in \mathcal{U} \quad \overrightarrow{OM}(u,v) = x(u,v)\overrightarrow{I} + y(u,v)\overrightarrow{J} + z(u,v)\overrightarrow{K}$$

L'application M est de classe C^k si et seulement si x, y et z le sont (ceci ne dépend évidemment pas du repère choisi).

Comme dans le cas des arcs paramétrés, on peut introduire une notion d'équivalence de nappes paramétrées :

DÉFINITION 21-1.2 *Soit $k \in \mathbb{N}^*$ et soient $\mathcal{U}$ et $\mathcal{V}$ deux ouverts connexes de $\mathbb{R}^2$. Deux nappes paramétrées*

$$S_1 : \mathcal{U} \ni (u,v) \mapsto M(u,v) \quad \text{et} \quad S_2 : \mathcal{V} \ni (s,t) \mapsto P(s,t)$$

sont C^k-équivalentes si et seulement s'il existe un C^k-difféomorphisme

$$\Phi : \mathcal{U} \to \mathcal{V}$$

tel que $M = P \circ \Phi$.

Φ est donc une bijection de classe C^k de $\mathcal{U}$ dans $\mathcal{V}$, dont la différentielle en tout point est un automorphisme de $\mathbb{R}^2$. Si

$$\forall (u,v) \in \mathcal{U} \quad \Phi(u,v) = (s(u,v), t(u,v))$$

on a donc

$$\forall (u,v) \in \mathcal{U} \quad \det J_\Phi(u,v) = \begin{vmatrix} \dfrac{\partial s}{\partial u}(u,v) & \dfrac{\partial s}{\partial v}(u,v) \\ \dfrac{\partial t}{\partial u}(u,v) & \dfrac{\partial t}{\partial v}(u,v) \end{vmatrix} \neq 0$$

et, pour tout $(u,v) \in \mathcal{U}$, $M(u,v) = P(s(u,v), t(u,v))$.

On définit ainsi une relation d'équivalence sur l'ensemble des nappes paramétrées de classe C^k. Une classe d'équivalence est une nappe géométrique, et une propriété géométrique d'une nappe est une propriété invariante par C^k-équivalence (on dit aussi par changement de paramétrage admissible). Par exemple, le support d'une nappe et la multiplicité d'un point M_0 du support

$$\{M(u,v),\ (u,v) \in \mathcal{U}\} \quad \text{et} \quad \operatorname{card}\{(u,v) \in \mathcal{U} \mid M(u,v) = M_0\}$$

sont des notions géométriques. Une nappe est dite simple si chaque point du support est de multiplicité 1, ce qui revient à dire que l'application M considérée est injective sur $\mathcal{U}$.

REMARQUE 21-1.3 Si $\Phi : \mathcal{U} \to \mathcal{V}$ est un C^k-difféomorphisme, l'application définie par $(u,v) \mapsto \det J_\Phi(u,v)$, continue sur le connexe $\mathcal{U}$, ne s'annule pas et garde donc un signe constant. On peut, comme dans le cas des arcs paramétrés, scinder l'ensemble des changements de paramétrage en deux classes

$$\mathcal{C}^+ = \{\Phi \mid \det J_\Phi > 0 \text{ sur } \mathcal{U}\} \quad \text{et} \quad \mathcal{C}^- = \{\Phi \mid \det J_\Phi < 0 \text{ sur } \mathcal{U}\}$$

A l'intérieur d'une même classe, deux paramétrages se déduisent l'un de l'autre par un difféomorphisme de déterminant Jacobien strictement positif. On définit ainsi une notion d'orientation sur une nappe géométrique. Nous verrons que, lorsque l'espace est euclidien et la nappe est simple et régulière, cela correspond à la notion intuitive d'orientation sur la normale à la nappe.

21-1.2 Plan tangent en un point régulier

DÉFINITION 21-1.4 *Soit $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ une nappe paramétrée de classe C^k. On appelle rang de la nappe en $(u_0,v_0) \in \mathcal{U}$ le rang de la différentielle de l'application M en (u_0,v_0). C'est donc le rang de la matrice Jacobienne*

$$\begin{pmatrix} \dfrac{\partial x}{\partial u}(u_0,v_0) & \dfrac{\partial x}{\partial v}(u_0,v_0) \\ \dfrac{\partial y}{\partial u}(u_0,v_0) & \dfrac{\partial y}{\partial v}(u_0,v_0) \\ \dfrac{\partial z}{\partial u}(u_0,v_0) & \dfrac{\partial z}{\partial v}(u_0,v_0) \end{pmatrix}$$

si $M(u,v) \left| \begin{array}{l} x(u,v) \\ y(u,v) \\ z(u,v) \end{array} \right.$ est représenté par ses coordonnées dans un repère affine $\mathcal{R} = \left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$. Comme on a

$$\begin{aligned} \frac{\partial \overrightarrow{M}}{\partial u}(u_0,v_0) &= \frac{\partial x}{\partial u}(u_0,v_0)\,\overrightarrow{I} + \frac{\partial y}{\partial u}(u_0,v_0)\,\overrightarrow{J} + \frac{\partial z}{\partial u}(u_0,v_0)\,\overrightarrow{K} \\ \frac{\partial \overrightarrow{M}}{\partial v}(u_0,v_0) &= \frac{\partial x}{\partial v}(u_0,v_0)\,\overrightarrow{I} + \frac{\partial y}{\partial v}(u_0,v_0)\,\overrightarrow{J} + \frac{\partial z}{\partial v}(u_0,v_0)\,\overrightarrow{K} \end{aligned}$$

ce rang est aussi égal au rang du système de vecteurs $\left(\dfrac{\partial \overrightarrow{M}}{\partial u}(u_0,v_0), \dfrac{\partial \overrightarrow{M}}{\partial v}(u_0,v_0)\right)$.

Il s'agit d'une notion qui se conserve par changement de paramétrage : si $\Phi : \mathcal{U} \to \mathcal{V}$ est un difféomorphisme et si $P : \mathcal{V} \to \mathcal{E}_3$ est telle que $M = P \circ \Phi$, on a par le théorème de différentiation d'une composée

$$dM_{(u_0,v_0)} = dP_{(s_0,t_0)} \circ d\Phi_{(u_0,v_0)}$$

si $(s_0,t_0) = \Phi(u_0,v_0)$. Comme $d\Phi_{(u_0,v_0)} \in \mathcal{GL}(\mathbb{R}^2)$, on a bien

$$\operatorname{rang} dM_{(u_0,v_0)} = \operatorname{rang} dP_{(s_0,t_0)}$$

Ce rang intervient de manière naturelle lorsqu'on cherche à déterminer les tangentes aux arcs tracés sur la nappe :

DÉFINITION 21-1.5 *Si $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ est une nappe de classe C^k, on appelle arc paramétré de classe C^k tracé sur la nappe toute application de la forme*

$$F = M \circ \varphi$$

où $\varphi : I \to \mathcal{U}$ est un arc de classe C^k dont le support est inclus dans $\mathcal{U}$ (I est un intervalle de $\mathbb{R}$ non réduit à un point).

Si φ est défini par $I \ni t \mapsto (u(t), v(t)) \in \mathcal{U}$ (les applications composantes étant de classe $\mathcal{C}^k$), nous aurons donc

$$\forall t \in I \quad F(t) = M(u(t), v(t))$$

Notations : par abus d'écriture, on utilisera souvent la notation $M(t)$ au lieu de $F(t)$; le point "générique" $M(u, v)$ devient une fonction du paramètre t.

Considérons un arc $\mathcal{C}^k$ tracé sur la nappe

$$\gamma : I \ni t \mapsto M(t) = M(u(t), v(t))$$

et supposons $0 \in I$ (on peut toujours s'y ramener par une translation sur le paramètre) et $(u_0, v_0) = (u(0), v(0))$. On a

$$\frac{d\overrightarrow{M}}{dt}(0) = u'(0)\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) + v'(0)\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$$

Si l'arc $I \ni t \mapsto (u(t), v(t))$ est régulier en 0, on a $(u'(0), v'(0)) \neq (0,0)$. Si la nappe est de rang 2 en (u_0, v_0), c'est à dire si les vecteurs $\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0)$ et $\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$ sont indépendants, le vecteur $\overrightarrow{M}'(0) = \frac{d\overrightarrow{M}}{dt}(0)$ est non nul et détermine la tangente à l'arc considéré en 0. On voit que ce vecteur appartient au plan vectoriel[1]

$$\Pi_{(u_0,v_0)} = \text{vect}\left(\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right)$$

De plus, tout vecteur non nul de ce plan vectoriel peut être obtenu comme vecteur dérivé à un arc tracé sur la surface : si $(\alpha, \beta) \in \mathbf{R}^2$ est non nul, il suffit de prendre $(u(t) = u_0 + \alpha t, v(t) = v_0 + \beta t)$ (point qui appartient bien à l'ouvert $\mathcal{U}$ pour $|t|$ "petit") pour avoir

$$\frac{d\overrightarrow{M}}{dt}(0) = \alpha\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) + \beta\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$$

On a donc :

[1] Si $(u'(0), v'(0)) = 0$, on a $\overrightarrow{M}'(0) = \overrightarrow{0}$. Si on suppose $k \geqslant 2$ et $(u''(0), v''(0)) \neq 0$, on aura

$$\frac{d^2\overrightarrow{M}}{dt^2}(t_0) = u''(0)\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) + v''(0)\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$$

puisque les autres termes apparaissant dans le calcul de dérivée seconde

$$u'^2(0)\frac{\partial^2 \overrightarrow{M}}{\partial u^2}(u_0, v_0) + 2u'(0)v'(0)\frac{\partial^2 \overrightarrow{M}}{\partial u \partial v}(u_0, v_0) + v'^2(0)\frac{\partial^2 \overrightarrow{M}}{\partial v^2}(u_0, v_0)$$

sont nuls. Le vecteur $\overrightarrow{M}''(0)$ est donc non nul, et c'est lui qui dirige la tangente à l'arc considéré. Ce vecteur est toujours dans le plan

$$\text{vect}\left(\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right)$$

On généraliserait ce résultat au cas où l'arc $t \mapsto (u(t), v(t))$ possède un premier indice fondamental en 0.

THÉORÈME 21-1.6 (et définition) Soit $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ une nappe paramétrée de classe $\mathcal{C}^k$. Un point $(u_0, v_0) \in \mathcal{U}$ est dit régulier pour cette nappe si le rang de la nappe en (u_0, v_0) est égal à 2, c'est à dire si

$$\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) \text{ et } \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0) \text{ sont indépendants}$$

Si (u_0, v_0) est régulier, la tangente en 0 à un arc régulier quelconque (défini au voisinage de 0) $t \mapsto M(u(t), v(t))$ tracé sur la nappe et vérifiant $(u(0), v(0)) = (u_0, v_0)$ est incluse dans le plan affine

$$T_{(u_0,v_0)} = M(u_0, v_0) + \operatorname{vect}\left(\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right)$$

Réciproquement, toute droite de ce plan passant par $M(u_0, v_0)$ est tangente à un tel arc. Ce plan est appelé plan tangent à la nappe en (u_0, v_0).

Ce plan est évidemment conservé par changement de paramétrage admissible.

DÉFINITION 21-1.7 *Une nappe paramétrée de classe C^k est régulière si et seulement si tous ses points sont réguliers.*

Dans un repère $\mathcal{R} = \left(O, \overrightarrow{\imath}, \overrightarrow{\jmath}, \overrightarrow{k}\right)$, où $M(u,v) \begin{vmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{vmatrix}$, un point $P \begin{vmatrix} X \\ Y \\ Z \end{vmatrix}$ appartient à ce plan si et seulement si

$$\det_{(\overrightarrow{\imath}, \overrightarrow{\jmath}, \overrightarrow{k})}\left(\overrightarrow{PM}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right) = 0$$

soit

$$\begin{vmatrix} X - x(u_0, v_0) & \dfrac{\partial x}{\partial u}(u_0, v_0) & \dfrac{\partial x}{\partial v}(u_0, v_0) \\ Y - y(u_0, v_0) & \dfrac{\partial y}{\partial u}(u_0, v_0) & \dfrac{\partial y}{\partial v}(u_0, v_0) \\ Z - z(u_0, v_0) & \dfrac{\partial z}{\partial u}(u_0, v_0) & \dfrac{\partial z}{\partial v}(u_0, v_0) \end{vmatrix} = 0$$

ce qui donne l'équation de $T_{(u_0,v_0)}$ dans le repère $\mathcal{R}$.

REMARQUE 21-1.8 Pour étudier la position d'un point $M(u,v)$ par rapport à $T_{(u_0,v_0)}$, on étudie le signe de la quantité

$$\delta(u,v) = \begin{vmatrix} x(u,v) - x(u_0, v_0) & \dfrac{\partial x}{\partial u}(u_0, v_0) & \dfrac{\partial x}{\partial v}(u_0, v_0) \\ y(u,v) - y(u_0, v_0) & \dfrac{\partial y}{\partial u}(u_0, v_0) & \dfrac{\partial y}{\partial v}(u_0, v_0) \\ z(u,v) - z(u_0, v_0) & \dfrac{\partial z}{\partial u}(u_0, v_0) & \dfrac{\partial z}{\partial v}(u_0, v_0) \end{vmatrix}$$

Il est clair que (u_0, v_0) est un point critique de δ. L'étude du signe de $\delta(u,v)$ au voisinage de (u_0, v_0) pourra donc se faire en étudiant la signature de la forme quadratique associée à la matrice Hessienne de δ en (u_0, v_0).

REMARQUE 21-1.9 Le point $(u_0, v_0) \in \mathcal{U}$ est dit singulier (ou stationnaire) si le rang de la nappe en (u_0, v_0) est strictement inférieur à 2 :

- Si rang $\left(\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0), \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right) = 1$ on peut supposer, sans nuire à la généralité, que $\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) \neq \overrightarrow{0}$ et qu'il existe un réel λ avec
$$\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0) = \lambda \frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0)$$
On a alors, avec les mêmes hypothèses que précédemment
$$\begin{aligned}\frac{d\overrightarrow{M}}{dt}(0) = u'(0)\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) + v'(0)\frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0) \\ = (u'(0) + \alpha v'(0))\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0)\end{aligned}$$
Tous les arcs réguliers tracés sur la nappe vérifiant $(u(0), v(0)) = (u_0, v_0)$ ont la même tangente en $t = 0$.
- Si $\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) = \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0) = \overrightarrow{0}$, tous les arcs tracés sur la nappe vérifiant $(u(0), v(0)) = (u_0, v_0)$ sont stationnaires en $t = 0$.

EXERCICE 21-1.10 Déterminer une nappe régulière $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ où $\mathcal{U}$ est un ouvert de $\mathbb{R}^2$ tel que le plan tangent en (u,v) ait pour équation
$$u^2x + v^2y + (1-u-v)^2 z = 0$$
dans un repère $\left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$. Cette nappe se prolonge de manière naturelle en une nappe dont le support contient l'origine O. Déterminer l'ensemble décrit par les tangentes en O aux arcs réguliers tracés sur cette nappe.

21-1.3 Normale, orientation

On suppose ici $\mathcal{E}_3$ euclidien orienté, rapporté à un repère orthonormé direct $\left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$ et une nappe paramétrée régulière $\mathcal{U} \ni (u,v) \mapsto M(u,v)$. En tout point $(u_0, v_0) \in \mathcal{U}$, le vecteur $\overrightarrow{N}(u_0, v_0)$ égal à $\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) \wedge \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$ est non nul, et est orthogonal au plan tangent à la nappe en (u_0, v_0) :

DÉFINITION 21-1.11 *Si $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ est une nappe régulière de classe C^k, le vecteur*
$$\overrightarrow{N}(u,v) = \frac{\partial \overrightarrow{M}}{\partial u}(u,v) \wedge \frac{\partial \overrightarrow{M}}{\partial v}(u,v)$$

est appelé vecteur normal à la nappe en $(u,v) \in \mathcal{U}$. *La droite*

$$\mathcal{N}_{(u,v)} = M(u,v) + \text{vect}\left(\overrightarrow{N}(u,v)\right)$$

est la normale à la nappe en (u,v).

Il est clair que, sous les hypothèses précédentes, l'application $(u,v) \mapsto \overrightarrow{N}(u,v)$ est de classe $\mathcal{C}^{k-1}$. L'équation du plan tangent en (u,v) peut aussi s'obtenir par la condition

$$P \in T_{u,v} \Leftrightarrow \overrightarrow{PM}(u,v).\overrightarrow{N}(u,v) = 0$$

Lorsqu'on effectue un changement de paramétrage admissible, on travaille avec une autre nappe régulière $\mathcal{V} \ni (s,t) \mapsto P(s,t)$ telle que $M = P \circ \Phi$, avec Φ $\mathcal{C}^k$-difféomorphisme de $\mathcal{U}$ dans $\mathcal{V}$, soit

$$M(u,v) = P(s(u,v), t(u,v))$$

On a alors, en $(u_0,v_0) \in \mathcal{U}$ correspondant à $(s_0,t_0) \in \mathcal{V}$,

$$\begin{cases} \dfrac{\partial \overrightarrow{M}}{\partial u}(u_0,v_0) = \dfrac{\partial s}{\partial u}(u_0,v_0)\dfrac{\partial \overrightarrow{P}}{\partial s}(s_0,t_0) + \dfrac{\partial t}{\partial u}(u_0,v_0)\dfrac{\partial \overrightarrow{P}}{\partial t}(s_0,t_0) \\ \dfrac{\partial \overrightarrow{M}}{\partial v}(u_0,v_0) = \dfrac{\partial s}{\partial v}(u_0,v_0)\dfrac{\partial \overrightarrow{P}}{\partial s}(s_0,t_0) + \dfrac{\partial t}{\partial v}(u_0,v_0)\dfrac{\partial \overrightarrow{P}}{\partial t}(s_0,t_0) \end{cases}$$

ce qui donne

$$\begin{aligned} &\frac{\partial \overrightarrow{M}}{\partial u}(u_0,v_0) \wedge \frac{\partial \overrightarrow{M}}{\partial v}(u_0,v_0) \\ &= \left(\frac{\partial s}{\partial u}(u_0,v_0)\frac{\partial t}{\partial v}(u_0,v_0) - \frac{\partial s}{\partial v}(u_0,v_0)\frac{\partial t}{\partial u}(u_0,v_0)\right)\frac{\partial \overrightarrow{P}}{\partial s}(s_0,t_0) \wedge \frac{\partial \overrightarrow{P}}{\partial t}(s_0,t_0) \end{aligned}$$

Les vecteurs normaux aux deux nappes sont évidemment colinéaires, et le coefficient de proportionnalité est det $J_\Phi(u_0,v_0)$, déterminant Jacobien du difféomorphisme Φ. On voit donc qu'un changement de paramétrage pour lequel ce déterminant est > 0 nous donne deux vecteurs normaux de même sens, alors que ceux-ci sont de sens opposés lorsque ce déterminant Jacobien est négatif. On comprend donc que le choix de l'une ou l'autre des classes de paramétrage $\mathcal{C}^+$ ou $\mathcal{C}^-$ envisagées à la remarque **21-1**.3 correspond (en général) au choix d'une orientation sur la normale à la nappe.

REMARQUE 21-1.12 Il faut savoir cependant qu'il existe des nappes

$$\mathcal{U} \ni (u,v) \mapsto M(u,v)$$

pour lesquelles il existe un $\mathcal{C}^k$ difféomorphisme Φ de $\mathcal{U}$ dans lui même, de déterminant Jacobien < 0 avec $M = M \circ \Phi$. Remplacer le couple (u,v) par $\Phi(u,v)$

donne alors le même point du support de la nappe, mais avec une normale orientée différemment (une telle nappe est dite "non orientable"). L'exemple le plus célèbre est sans doute le ruban de Möbius, défini en repère orthonormé par

$$\left]-\frac{\pi}{2},\frac{\pi}{2}\right[\times \mathbb{R} \ni (u,v) \mapsto M(u,v) \left| \begin{array}{l} R\,(2+\sin u\cos v)\cos 2v \\ R\,(2+\sin u\cos v)\sin 2v \\ R\,\sin u\sin v \end{array} \right.$$

où $R > 0$ est fixé. On vérifie que $\Phi(u,v) = (-u, v+\pi)$ a un déterminant Jacobien égal à -1, et vérifie $M \circ \Phi = M$. En prenant par exemple $u = 0$, et en faisant varier v entre 0 et π, un observateur placé sur la nappe décrit un cercle. Il revient à son point de départ avec la tête en bas.

21-1.4 Exemples

21-1.4.1 Nappe définie par $z = f(x,y)$

On considère $\mathcal{U}$ un ouvert connexe de $\mathbb{R}^2$ et $f : \mathcal{U} \to \mathbb{R}$ de classe C^k. La nappe étudiée ici est parfois appelée "nappe cartésienne", elle est définie dans un repère $\left(O, \vec{I}, \vec{J}, \vec{K}\right)$ par

$$\mathcal{U} \ni (x,y) \mapsto M(x,y) \left| \begin{array}{c} x \\ y \\ f(x,y) \end{array} \right.$$

Les résultats des sections précédentes correspondent à ceux obtenus au chapitre sur les fonctions de plusieurs variables (section **18-2.3.2**) : une telle nappe est évidemment simple et régulière, puisque

$$\frac{\partial \vec{M}}{\partial x}(x,y) = \vec{I} + \frac{\partial f}{\partial x}(x,y)\,\vec{K} \quad \text{et} \quad \frac{\partial \vec{M}}{\partial y}(x,y) = \vec{J} + \frac{\partial f}{\partial y}(x,y)\,\vec{K}$$

En $(x_0, y_0) \in \mathcal{U}$, l'équation du plan tangent est

$$P \left| \begin{array}{c} X \\ Y \\ Z \end{array} \right. \in T_{(x_0,y_0)} \Leftrightarrow \begin{vmatrix} X - x_0 & 1 & 0 \\ Y - y_0 & 0 & 1 \\ Z - f(x_0,y_0) & \dfrac{\partial f}{\partial x}(x_0,y_0) & \dfrac{\partial f}{\partial y}(x_0,y_0) \end{vmatrix} = 0$$

ce qui peut s'écrire

$$Z - f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)\,(X - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\,(Y - y_0)$$

Rappelons que, lorsque $k \geqslant 2$, la position locale du support de la nappe par rapport à son plan tangent s'étudie en fonction de la signature de la forme quadratique associée à la matrice hessienne de f (voir proposition **18-3**.22).

Nous verrons ultérieurement que (localement) une nappe paramétrée régulière admet un paramétrage admissible correspondant (avec un bon choix du repère) à une nappe cartésienne.

21-1.4.2 Nappe cylindrique

DÉFINITION 21-1.13 *Soient Γ une "courbe" de l'espace $\mathcal{E}_3$ et $\overrightarrow{u}$ un vecteur non nul. Le cylindre de directrice Γ et de génératrices parallèles à $\overrightarrow{u}$ est la réunion des droites dirigées par $\overrightarrow{u}$ et rencontrant Γ. Toute droite de la forme $P_0 + \mathbb{R}\overrightarrow{u}$ avec $P_0 \in \Gamma$ est appelée génératrice du cylindre.*

Si C désigne ce cylindre, on a donc, pour $M \in \mathcal{E}_3$,

$$M \in C \Leftrightarrow \exists P_0 \in \Gamma \quad \exists s \in \mathbb{R} \quad \overrightarrow{P_0M} = s\overrightarrow{u}$$

Si Γ est le support d'un arc paramétré γ de classe C^k

$$\gamma : \left]a, b\right[\ni t \mapsto P(t)$$

on peut interpréter C comme support d'une nappe paramétrée de classe C^k définie sur l'ouvert $\mathcal{U} = \mathbb{R} \times \left]a, b\right[$ par

$$\mathbb{R} \times \left]a, b\right[\ni (s, t) \mapsto M(s, t) = P(t) + s\overrightarrow{u}$$

Dans un repère $\left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$ où $P(t) \left| \begin{array}{l} f(t) \\ g(t) \\ h(t) \end{array} \right.$ et $\overrightarrow{u} \left| \begin{array}{l} \alpha \\ \beta \\ \gamma \end{array} \right.$, cette nappe nous fournit une "représentation paramétrique" de C :

$$M(s, t) \left| \begin{array}{l} f(t) + \alpha s \\ g(t) + \beta s \\ h(t) + \gamma s \end{array} \right.$$

On a évidemment, pour $(s, t) \in \mathbb{R} \times \left]a, b\right[$

$$\frac{\partial \overrightarrow{M}}{\partial t}(s, t) = \overrightarrow{P}'(t) \quad \text{et} \quad \frac{\partial \overrightarrow{M}}{\partial s}(s, t) = \overrightarrow{u}$$

- En général, le rang de $\left(\overrightarrow{P}'(t), \overrightarrow{u}\right)$ égal à 2. Le point (s, t) est donc régulier, et le plan tangent en ce point est

  $$T_{(s,t)} = M(s, t) + \text{vect}\left(\overrightarrow{P}'(t), \overrightarrow{u}\right) = P(t) + \text{vect}\left(\overrightarrow{P}'(t), \overrightarrow{u}\right)$$

  Ce plan tangent est donc constant le long de la génératice contenant $M(s, t)$. C'est le plan déterminé par la génératrice et la tangente à γ en t.

- Le point (s, t) est singulier lorsque $\overrightarrow{P}'(t) = \overrightarrow{0}$ (t est un point stationnaire de γ) ou lorsque le vecteur $\overrightarrow{P}'(t)$ est non nul et colinéaire à $\overrightarrow{u}$. Dans ce cas, la génératrice du cylindre contenant $M(s, t)$ se confond avec la tangente à l'arc γ en t (et tous les points de cette génératrice correspondent à des points stationnaires de la nappe).

Pour déterminer un cylindre, on en cherche souvent une section plane (intersection avec un plan non parallèle à $\overrightarrow{u}$, qui contient un et un seul point de chaque génératrice). Lorsque l'espace est euclidien et le plan orthogonal à $\overrightarrow{u}$, on dit que la section est droite.

21-1.4.3 Nappe conique

DÉFINITION 21-1.14 *Soient Γ une "courbe" de l'espace $\mathcal{E}_3$ et $S \in \mathcal{E}_3 - \Gamma$ un point quelconque. Le cône de directrice Γ et de sommet S est la réunion des droites passant par S et rencontrant Γ. Toute droite de la forme $S + \mathbb{R}\overrightarrow{SP}_0$ avec $P_0 \in \Gamma$ est appelée génératrice du cône.*

Si $\mathcal{C}$ désigne ce cône, on a alors

$$M \in \mathcal{C} \Leftrightarrow \exists P_0 \in \Gamma \quad \exists s \in \mathbb{R} \quad \overrightarrow{SM} = s\,\overrightarrow{SP}_0$$

Ainsi, si Γ est le support d'un arc paramétré γ de classe $\mathcal{C}^k$

$$\gamma : \,]a,b[\, \ni t \mapsto P(t)$$

on peut interpréter $\mathcal{C}$ comme support d'une nappe paramétrée de classe $\mathcal{C}^k$ définie sur l'ouvert $\mathcal{U} = \mathbb{R} \times \,]a,b[$ par

$$\mathbb{R} \times \,]a,b[\, \ni (s,t) \mapsto M(s,t) = S + s\,\overrightarrow{SP}(t)$$

Dans un repère $\left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$ où $P(t) \left| \begin{array}{l} f(t) \\ g(t) \\ h(t) \end{array} \right.$ et $S \left| \begin{array}{l} \alpha \\ \beta \\ \gamma \end{array} \right.$, on obtient une représentation paramétrique de $\mathcal{C}$:

$$M(s,t) \left| \begin{array}{l} s\,f(t) + (1-s)\,\alpha \\ s\,g(t) + (1-s)\,\beta \\ s\,h(t) + (1-s)\,\gamma \end{array} \right.$$

On a ici, pour $(s,t) \in \mathbb{R} \times \,]a,b[$

$$\frac{\partial \overrightarrow{M}}{\partial t}(s,t) = s\,\overrightarrow{P}'(t) \quad \text{et} \quad \frac{\partial \overrightarrow{M}}{\partial s}(s,t) = \overrightarrow{SP}(t)$$

- Pour $s \neq 0$ et $\left(\overrightarrow{P}'(t), \overrightarrow{SP}(t)\right)$ indépendants, le point (s,t) est régulier, et le plan tangent en ce point est

  $$T_{(s,t)} = M(s,t) + \mathrm{vect}\left(\overrightarrow{P}'(t), \overrightarrow{SP}(t)\right) = S + \mathrm{vect}\left(\overrightarrow{P}'(t), \overrightarrow{SP}(t)\right)$$

  Ce plan tangent est donc constant le long de la génératice contenant $M(s,t)$. C'est le plan déterminé par la génératrice et la tangente à γ en t.

- Lorsque $\overrightarrow{P}'(t) = \overrightarrow{0}$ (t est un point stationnaire de γ) ou $s = 0$ (on a alors $M(s,t) = S$), le point (s,t) est singulier. C'est aussi le cas lorsque le vecteur $\overrightarrow{P}'(t)$ est non nul et colinéaire à $\overrightarrow{SP}(t)$. Dans ce cas, la génératrice du cylindre contenant $M(s,t)$ se confond avec la tangente à l'arc γ en t (et tous les points de cette génératrice correspondent à des points stationnaires de la nappe).

Si un plan $\mathcal{P}$ ne passant pas par S rencontre toutes les génératrices d'un cône $\mathcal{C}$, ce dernier est entièrement déterminé par la donnée de S et de $\mathcal{C} \cap \mathcal{P}$. Cette intersection est alors appelée base du cône.

21-1.4.4 Nappe de révolution

L'espace $\mathcal{E}_3$ est évidemment muni ici d'une structure euclidienne.

DÉFINITION 21-1.15 *Soit Γ une courbe de $\mathcal{E}_3$ et $\Delta \subset \mathcal{E}_3$ une droite. La surface de révolution $\mathcal{S}$ de directrice Γ et d'axe Δ est la réunion des cercles d'axe Δ rencontrant Γ. Un tel cercle est un parallèle de $\mathcal{S}$. On appelle plan méridien (pour $\mathcal{S}$) tout plan contenant Δ. L'intersection de $\mathcal{S}$ avec un plan méridien quelconque est une méridienne de $\mathcal{S}$.*

$\mathcal{S}$ est donc la surface engendrée par la rotation de Γ autour de Δ. Si $\mathcal{P}$ est un plan méridien, tout parallèle rencontre $\mathcal{P}$ en deux points (symétriques par rapport à Δ, éventuellement confondus si ce cercle se réduit à un point de Δ). La surface $\mathcal{S}$ est donc aussi engendré par la rotation d'une des ses méridiennes autour de Δ.

Pour décrire $\mathcal{S}$ on utilise souvent un repère $\left(O, \overrightarrow{I}, \overrightarrow{J}, \overrightarrow{K}\right)$ pour lequel la droite Δ se confond avec l'axe $O + \mathbb{R}\overrightarrow{K}$:

- Si Γ est le support d'un arc paramétré de classe $\mathcal{C}^k$ s'exprimant dans ce repère par

$$\gamma :]a, b[\ni t \mapsto P(t) \left| \begin{array}{l} f(t) \\ g(t) \\ h(t) \end{array} \right.$$

un point $M \left| \begin{array}{l} x \\ y \\ z \end{array} \right. \in \mathcal{E}_3$ appartient à $\mathcal{S}$ si et seulement s'il existe $t \in]a, b[$ et $\theta \in \mathbb{R}$ tel que M se déduise de $P(t)$ par la rotation d'axe Oz et d'angle θ. On a donc

$$M \left| \begin{array}{l} x \\ y \\ z \end{array} \right. \in \mathcal{S} \Leftrightarrow \exists t \in]a, b[\quad \exists \theta \in \mathbb{R} \quad \left\{ \begin{array}{l} x = f(t)\cos\theta - g(t)\sin\theta \\ y = f(t)\sin\theta + g(t)\cos\theta \\ z = h(t) \end{array} \right.$$

La surface $\mathcal{S}$ peut donc être interprétée comme support d'une nappe paramétrée de classe $\mathcal{C}^k$

$$]a, b[\times \mathbb{R} \ni (t, \theta) \mapsto M(t, \theta)$$
$$\text{avec} \quad \overrightarrow{OM}(t, \theta) = f(t)\,\overrightarrow{u}(\theta) + g(t)\,\overrightarrow{v}(\theta) + h(t)\,\overrightarrow{K}$$

avec les notations usuelles des coordonnées cylindriques d'axe Oz.

- Si la méridienne dans le plan de coordonnées $O + \text{vect}\left(\overrightarrow{I}, \overrightarrow{K}\right)$ est le support d'un arc $\mathcal{C}^k$

$$]a, b[\ni t \mapsto P_1(t) \left| \begin{array}{l} x(t) = \rho(t) \\ z(t) = h(t) \end{array} \right.$$

on aura plus simplement une représentation de $\mathcal{S}$:

$$]a, b[\times \mathbb{R} \ni (t, \theta) \mapsto M(t, \theta) \quad \text{avec} \quad \overrightarrow{OM}(t, \theta) = \rho(t)\,\overrightarrow{u}(\theta) + h(t)\,\overrightarrow{K}$$

Sous cette forme, on a évidemment

$$\frac{\partial \overrightarrow{M}}{\partial t}(t,\theta) = \rho'(t)\,\overrightarrow{u}(\theta) + h'(t)\,\overrightarrow{K} \quad \text{et} \quad \frac{\partial \overrightarrow{M}}{\partial \theta}(t,\theta) = \rho(t)\,\overrightarrow{v}(\theta)$$

On en déduit

$$\overrightarrow{N}(t,\theta) = \frac{\partial \overrightarrow{M}}{\partial t}(t,\theta) \wedge \frac{\partial \overrightarrow{M}}{\partial \theta}(t,\theta) = \rho(t)\left[-h'(t)\,\overrightarrow{u}(\theta) + \rho'(t)\,\overrightarrow{K}\right]$$

Le point (t,θ) est donc stationnaire si $\rho(t) = 0$ (le point correspondant de la méridienne est sur l'axe Δ) ou si $\rho'(t) = h'(t) = 0$ (le point t est stationnaire pour la méridienne). Hormis ces cas particuliers, le vecteur normal en (t,θ) est colinéaire au vecteur $\overrightarrow{n_\theta}(t) = -h'(t)\,\overrightarrow{u}(\theta) + \rho'(t)\,\overrightarrow{K}$, qu'on peut interpréter comme normal[2] en t à la méridienne contenant $M(t,\theta)$.

21-1.5 Définition par paramétrage et par équation

21-1.5.1 Equivalence locale

- Si $\mathcal{E}_3$ est rapporté à un repère $\left(O,\overrightarrow{I},\overrightarrow{J},\overrightarrow{K}\right)$ et si f est une fonction de classe C^k ($k \geqslant 1$) définie sur un ouvert U de $\mathbb{R}^3$, la "surface" d'équation $f(x,y,z) = 0$ est, par définition

$$\Sigma_f = \{M(x,y,z) \in \mathcal{E}_3 \mid (x,y,z) \in U \quad \text{et} \quad f(x,y,z) = 0\}$$

Rappelons qu'un point $M_0 \in \Sigma_f$ est dit **régulier** si et seulement si $df_{M_0} \neq 0$. Le théorème des fonctions implicites (section **18-3.4.3**) nous a montré qu'au voisinage d'un point régulier M_0, Σ_f peut être considéré comme support d'une nappe paramétrée de classe C^k. Plus précisément, si $M(x_0,y_0,z_0)$ appartient à Σ_f et vérifie $\frac{\partial f}{\partial z}(M_0) \neq 0$, il existe une fonction φ de classe C^k définie sur un voisinage ouvert $\mathcal{U}$ de (x_0,y_0) à valeur dans un voisinage ouvert $\mathcal{V}$ de z_0

$$\mathcal{U} \ni (x,y) \mapsto \varphi(x,y) \in \mathcal{V}$$

telle que

$$\Sigma_f \cap (\mathcal{U} \times \mathcal{V}) = \{M(x,y,z) \mid (x,y) \in \mathcal{U} \text{ et } z = \varphi(x,y)\}$$

$\Sigma_f \cap (\mathcal{U} \times \mathcal{V})$ est donc le support de la nappe paramétrée (cartésienne) définie par $\mathcal{U} \ni (x,y) \mapsto M(x,y,\varphi(x,y))$. Une telle nappe est **régulière**, au sens de la définition donnée à la section **21-1.2**.

[2]Le vecteur

$$\frac{\partial \overrightarrow{M}}{\partial t}(t,\theta) = \rho'(t)\,\overrightarrow{u}(\theta) + h'(t)\,\overrightarrow{K}$$

dirige la tangente en $M(t,\theta)$ à cette méridienne, alors que

$$\frac{\partial \overrightarrow{M}}{\partial \theta}(t,\theta) = \rho(t)\,\overrightarrow{v}(\theta)$$

dirige la tangente en $M(t,\theta)$ au parallèle passant par ce point.

- Réciproquement, soit $\mathcal{U} \ni (u,v) \mapsto M(u,v)$ une nappe paramétrée. En travaillant dans un repère $\left(O, \vec{I}, \vec{J}, \vec{K}\right)$, on a

$$\overrightarrow{OM}(u,v) = f(u,v)\,\vec{I} + g(u,v)\,\vec{J} + h(u,v)\,\vec{K}$$

Si (u_0, v_0) est un point régulier de cette nappe, la matrice

$$\begin{pmatrix} \dfrac{\partial f}{\partial u}(u_0,v_0) & \dfrac{\partial g}{\partial u}(u_0,v_0) & \dfrac{\partial h}{\partial u}(u_0,v_0) \\ \dfrac{\partial f}{\partial v}(u_0,v_0) & \dfrac{\partial g}{\partial v}(u_0,v_0) & \dfrac{\partial h}{\partial v}(u_0,v_0) \end{pmatrix}$$

est de rang 2 et, sans nuire à la généralité, on peut supposer que le déterminant formé par ses deux premières colonnes est non nul

$$\begin{vmatrix} \dfrac{\partial f}{\partial u}(u_0,v_0) & \dfrac{\partial g}{\partial u}(u_0,v_0) \\ \dfrac{\partial f}{\partial v}(u_0,v_0) & \dfrac{\partial g}{\partial v}(u_0,v_0) \end{vmatrix} \neq 0$$

Le théorème d'inversion locale montre alors qu'il existe un voisinage ouvert $\mathcal{U}_0 \subset U$ de (u_0, v_0) et un voisinage ouvert $\mathcal{V}_0$ de

$$m_0 = (x_0 = f(u_0,v_0), y_0 = g(u_0,v_0))$$

dans $\mathbb{R}^2$ tels que l'application induite

$$\Phi : \mathcal{U}_0 \to \mathcal{V}_0 \quad (u,v) \mapsto (f(u,v), g(u,v))$$

soit un C^k-difféomorphisme. On peut considérer que cette application définit, au voisinage de (u_0, v_0), un changement de paramétrage admissible. Si nous notons Ψ le difféomorphisme réciproque

$$\Psi : \mathcal{V}_0 \to \mathcal{U}_0 \quad (x,y) \mapsto (u = \psi_1(x,y), v = \psi_2(x,y))$$

la nappe définie par

$$\mathcal{V}_0 \ni (x,y) \mapsto P(x,y) = M \circ \Psi(y,x) \left| \begin{array}{c} x \\ y \\ h(\psi_1(x,y), \psi_2(x,y)) \end{array} \right.$$

est équivalente à $M|_{\mathcal{U}_0}$. Cette dernière est donc simple, et son support est exactement

$$\{N(x,y,z) \in \mathcal{E}_3 \mid (x,y) \in \mathcal{V}_0 \text{ et } z = h(\psi_1(x,y), \psi_2(x,y))\}$$

Ce support est donc défini par une équation cartésienne de la forme

$$F(x,y,z) = 0$$

avec F de classe C^k définie sur $\mathcal{V}_0$ par

$$F(x,y,z) = z - h(\psi_1(x,y), \psi_2(x,y))$$

Il y a donc, **localement**, et sous une hypothèse de **régularité**, équivalence entre la définition d'une surface par une équation ou comme support d'une nappe paramétrée.

Le passage de la représentation paramétrique

$$\mathcal{U} \ni (u,v) \mapsto M\left(f(u,v), g(u,v), h(u,v)\right)$$

à la représentation à l'aide d'une équation est un problème d'**élimination** : si $(x,y,z) \in \mathbb{R}^3$, il s'agit de trouver une condition nécessaire (et on l'espère suffisante) pour que le système d'équations

$$\begin{cases} x = f(u,v) \\ y = g(u,v) \quad \text{avec } (u,v) \in \mathcal{U} \\ z = h(u,v) \end{cases}$$

ait au moins une solution.

Théoriquement, le théorème des fonctions implicites permet de passer d'une représentation par équation à un paramétrage. On ne peut trouver **explicitement** un paramétrage en partant d'une équation que dans des cas tout à fait particuliers.

21-1.5.2 Exemples : cylindres, cônes et surfaces de révolution

CYLINDRES :

Par cylindre, nous entendons ici réunion d'une famille de droites parallèles à une direction donnée. Considérons un cylindre $\mathcal{C}$ de génératrices parallèles à un vecteur $\overrightarrow{u}$, caractérisé par une section plane $\mathcal{C} \cap \mathcal{P}$, où $\mathcal{P}$ est un plan dont la direction ne contient pas $\overrightarrow{u}$. On choisit un repère affine d'origine Ω pour lequel le plan $\mathcal{P}$ est le plan $X\Omega Y$ d'équation $Z = 0$, l'axe ΩZ étant dirigé par $\overrightarrow{u}$. Si $\mathcal{C} \cap \mathcal{P}$ admet une équation cartésienne

$$M(X,Y,0) \in \mathcal{C} \cap \mathcal{P} \Leftrightarrow F(X,Y) = 0$$

on obtiendra une équation de $\mathcal{C}$ en remarquant que, les génératrices du cylindre étant parallèles à ΩZ,

$$M(X,Y,Z) \in \mathcal{C} \Leftrightarrow M_0(X,Y,0) \in \mathcal{C} \cap \mathcal{P}$$

L'équation $F(X,Y) = 0$ caractérise donc les points de $\mathcal{C}$. Réciproquement, il est clair que l'ensemble des points vérifiant une telle équation est invariant par translation parallèlement à OZ, et est donc un cylindre de génératrices parallèles à OZ (axe dont un système d'équations est $X = Y = 0$).

Si on travaille à présent dans un repère quelconque $\left(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\right)$, où les coordonnées du point "générique" sont notées (x,y,z), les formules de changement de repères nous donnent des relations de la forme

$$\begin{cases} X = \alpha x + \beta y + \gamma z + \delta \\ Y = \alpha' x + \beta' y + \gamma' z + \delta' \end{cases}$$

Les coordonnées X et Y sont des formes affines indépendantes[3] de (x,y,z). Nous avons donc :

[3] Cela signifie que les formes linéaires associées, donc ici les lignes (α,β,γ) et (α',β',γ'), sont indépendantes.

PROPOSITION 21-1.16 L'espace $\mathcal{E}_3$ est rapporté à un repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$. Si

$$P_1(x,y,z) = 0 \quad \text{et} \quad P_2(x,y,z) = 0$$

sont deux équations de plans $\mathcal{P}_1$ et $\mathcal{P}_2$ non parallèles (les formes affines P_1 et P_2 sont indépendantes), et si F est une fonction numérique définie sur une partie de $\mathbb{R}^2$, l'ensemble des points $M(x,y,z)$ de $\mathcal{E}_3$ vérifiant l'équation

$$F(P_1(x,y,z), P_2(x,y,z)) = 0$$

est un cylindre (éventuellement $\emptyset$) dont les génératrices sont parallèles à la droite

$$\mathcal{D} = \mathcal{P}_1 \cap \mathcal{P}_2$$

Une équation de cylindre est donc, d'une certaine manière, "dégénérée" : dans un bon repère, une des coordonnées n'apparaît pas.

EXEMPLE 21-1.17 Dans $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$, l'ensemble des points vérifiant l'équation

$$x^2 - y^2 - z^2 + 2yz - 3y - 3z = 0$$

est un cylindre dont les génératrices sont parallèles au vecteur $\vec{u} = \vec{\jmath} - \vec{k}$. En effet, l'équation peut s'écrire

$$x^2 - (y-z)^2 - 3(y-z) = 0$$

et fait intervenir les formes affines indépendantes

$$P_1(x,y,z) = x \text{ et } P_2(x,y,z) = y - z$$

Ce cylindre possède une section plane qui est une hyperbole : si on complète ces deux formes par une troisième (indépendante des deux premières) par exemple $P_3(x,y,z) = y + z$, on aura l'équation

$$X^2 - Y^2 - 3Y = 0$$

dans un repère où les coordonnées seraient données par

$$X = P_1(x,y,z),\ Y = P_2(x,y,z) \text{ et } Z = P_3(x,y,z)$$

L'intersection avec le plan de coordonnées $Z = 0$ est une courbe du second degré qui est bien du genre hyperbole.

EXERCICE 21-1.18 Soit une surface $\mathcal{S}$ d'équation

$$f(x,y,z) = 0$$

dans un repère $\left(O, \vec{\imath}, \vec{\jmath}, \vec{k}\right)$, où f est une fonction de classe $\mathcal{C}^k$ ($k \geqslant 1$), tous les points de $\mathcal{S}$ étant supposés réguliers. Si

$$\vec{u} = \alpha\vec{\imath} + \beta\vec{\jmath} + \gamma\vec{k}$$

est un vecteur non nul, le cylindre de génératrices parallèles à $\overrightarrow{u}$ circonscrit à $\mathcal{S}$ est, par définition, la réunion des droites dirigées par $\overrightarrow{u}$ et tangentes à $\mathcal{S}$. Montrer qu'on peut trouver une équation de ce cylindre en éliminant le paramètre λ dans le système d'équations

$$\begin{cases} f(x+\lambda\alpha, y+\lambda\beta, z+\lambda\gamma) = 0 \\ \alpha\dfrac{\partial f}{\partial x}(x+\lambda\alpha, y+\lambda\beta, z+\lambda\gamma) + \beta\dfrac{\partial f}{\partial y}(x+\lambda\alpha, y+\lambda\beta, z+\lambda\gamma) \\ \qquad\qquad + \gamma\dfrac{\partial f}{\partial z}(x+\lambda\alpha, y+\lambda\beta, z+\lambda\gamma) = 0 \end{cases}$$

Trouver de même un système d'équations du contour apparent de $\mathcal{S}$ dans la direction vect $(\overrightarrow{u})$, intersection de ce cylindre avec la surface.
Déterminer par exemple une équation du cylindre circonscrit au paraboloïde d'équation

$$(x^2+y^2) = 2pz$$

dans la direction du vecteur $\overrightarrow{u}(\alpha,\beta,\gamma)$ avec $(\alpha,\beta) \neq (0,0)$. Réponse :

$$(\beta x - \alpha y)^2 + 2p\left(\alpha\gamma x + \beta\gamma y - \left(\alpha^2+\beta^2\right)z\right) - p^2\gamma^2 = 0$$

CÔNES :

Les termes "cône de sommet Ω" désignent ici une réunion d'une famille de droites[4] passant par le point Ω. Soit $\mathcal{U}$ une partie de $\mathbb{R}^3$ stable par les homothéties vectorielles :

$$\forall\,(x_0,y_0,z_0)\in\mathcal{U} \quad \forall\,\lambda\in\mathbb{R} \quad \lambda\,(x_0,y_0,z_0)\in\mathcal{U}$$

Une fonction $F:\mathcal{U}\to\mathbb{R}$ est homogène de degré $p\in\mathbb{N}$ si et seulement si

$$\forall\,(x_0,y_0,z_0)\in\mathcal{U} \quad \forall\,\lambda\in\mathbb{R} \quad F(\lambda x_0,\lambda y_0,\lambda z_0) = \lambda^p F(x_0,y_0,z_0)$$

Soit un repère d'origine Ω, où les coordonnées du point générique sont notées (X,Y,Z). Si F est homogène de degré p, l'ensemble

$$\mathcal{C} = \{M(X,Y,Z) \mid F(X,Y,Z) = 0\}$$

est (s'il est non vide et non réduit à $\{\Omega\}$) un cône de sommet Ω : si $M_0(X_0,Y_0,Z_0)$ est un point de $\mathcal{C}$ distinct de Ω

$$F(X_0,Y_0,Z_0) = 0 \Rightarrow \forall\,\lambda\in\mathbb{R} \quad F(\lambda X_0,\lambda Y_0,\lambda Z_0) = 0$$

et il est clair que la droite ΩM_0 est incluse dans $\mathcal{C}$.

En effectuant un changement de repère comme dans le cas des cylindres, nous obtenons :

[4] On pourrait aussi étudier des demi-droites ouvertes issues de Ω. L'étude de ce paragraphe pourrait être menée de manière analogue, en introduisant la notion de fonction positivement homogène, vérifiant

$$\forall\,(x_0,y_0,z_0)\in\mathcal{U} \quad \forall\,\lambda\in\mathbb{R}^{*+} \quad F(\lambda x_0,\lambda y_0,\lambda z_0) = \lambda^\alpha F(x_0,y_0,z_0)$$

où α est à présent un réel quelconque. Voir à ce sujet la section **18-3.5**.

PROPOSITION 21-1.19 L'espace $\mathcal{E}_3$ est rapporté à un répère $\left(O, \vec{I}, \vec{J}, \vec{K}\right)$. Soient

$$P_1(x,y,z) = 0, \ P_2(x,y,z) = 0 \text{ et } P_2(x,y,z) = 0$$

trois équations de plans $\mathcal{P}_1$, $\mathcal{P}_2$ et $\mathcal{P}_3$ dont l'intersection est réduite à un point

$$\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3 = \{\Omega\}$$

(les formes affines P_1, P_2 et P_3 sont donc indépendantes). Si F est une fonction numérique homogène de degré $p \in \mathbb{N}$ définie sur une partie de $\mathbf{R}^3$, l'ensemble des points $M(x,y,z)$ de $\mathcal{E}_3$ vérifiant l'équation

$$F(P_1(x,y,z), P_2(x,y,z), P_3(x,y,z)) = 0$$

est un cône (éventuellement vide ou réduit à $\{\Omega\}$) de sommet Ω.

EXEMPLE 21-1.20 L'équation

$$(x-2)^2 - (x+y-5)(x-z+2) = 0$$

définit un cône de sommet Ω, dont les coordonnées vérifient

$$x - 2 = 0, \ x + y - 5 = 0 \text{ et } x - z + 2 = 0$$

soit $\Omega(2,3,4)$.

EXERCICE 21-1.21 L'espace euclidien $\mathcal{E}_3$ est rapporté à un repère orthonormé. Si $a > 0$, déterminer une équation du cône de sommet $\Omega(a,a,a)$ et dont une directrice est le cercle passant par les points $A(a,0,0)$, $B(0,a,0)$ et $C(0,0,a)$.

EXERCICE 21-1.22 Comme à l'exercice **21-1.18**, définir le cône de sommet Ω circonscrit à une surface $\mathcal{S}$ d'équation $F(x,y,z) = 0$ (on suppose $\Omega \notin \mathcal{S}$) et le contour apparent de la surface $\mathcal{S}$ vue de Ω. Dire comment en trouver des équations.

EXERCICE 21-1.23 Soit $\varphi : \mathcal{U} \to \mathbf{R}$ une fonction de classe $\mathcal{C}^1$, avec $\mathcal{U}$ ouvert de $\mathbf{R}^2$ stable par les homothéties de rapport strictement positif. On considère la nappe cartésienne définie par l'équation $z = \varphi(x,y)$. Montrer que le support $\mathcal{C}$ de cette nappe est inclus dans un cône de sommet O si et seulement si

$$\forall M \in \mathcal{C} \quad \overrightarrow{OM}.\vec{N}(M) = 0$$

où $\vec{N}(M)$ désigne le vecteur normal à la nappe au point M (on pourra utiliser les résultats de la section **18-3.5**).

SURFACES DE REVOLUTION :

Une surface de révolution d'axe Δ est ici considérée comme réunion d'une famille de cercles d'axe Δ. Soit $\mathcal{S}$ une telle surface et Ω un point choisi arbitrairement sur Δ. Dans un repère **orthonormé** d'origine Ω, et d'axe ΩZ égal à Δ, un cercle d'axe Δ peut être considéré comme intersection d'une sphère de centre

Ω et d'un plan parallèle à $X\Omega Y$. Il possède donc un système d'équations de la forme

$$\begin{cases} X^2 + Y^2 + Z^2 = \alpha^2 \\ Z = \beta \end{cases}$$

Il en résulte que toute équation de la forme

$$F\left(X^2 + Y^2 + Z^2, Z\right) = 0 \quad (*)$$

où F est une fonction définie sur une partie de $\mathbb{R}^2$ définit une surface de révolution $\mathcal{S}$ d'axe ΩZ (éventuellement vide ou réduite à une partie de l'axe ΩZ) : si $M_0\,(X_0, Y_0, Z_0)$ vérifie cette équation, tout point $M\,(X, Y, Z)$ appartenant au cercle engendré par la rotation de M_0 autour de ΩZ la vérifie également, puisqu'alors on a

$$Z = Z_0 \quad \text{et} \quad X^2 + Y^2 + Z^2 = X_0^2 + Y_0^2 + Z_0^2$$

En définissant la fonction G sur une partie convenable de $\mathbb{R}^2$ par

$$G\,(U, V) = F\left(U + V^2, V\right)$$

on préfère souvent écrire $(*)$ sous la forme

$$G\left(X^2 + Y^2, Z\right) = 0 \quad (**)$$

L'interprétation géométrique de cette équation est simple : $X^2 + Y^2$ représente le carré de la distance du point $M\,(X, Y, Z)$ à l'axe ΩZ. Elle se conserve, comme Z, par rotation autour de cet axe. Sous cette forme, l'équation $G\,(X^2, Z) = 0$ définit la méridienne de la surface dans le plan $X\Omega Z$.

Par changement de repère, l'écriture d'une équation de $\mathcal{S}$ sous la forme $(*)$ nous donne :

PROPOSITION 21-1.24 Soit Ω un point de l'espace euclidien $\mathcal{E}_3$ et soit $M \mapsto P\,(M)$ une forme affine définie sur $\mathcal{E}_3$, l'équation $P\,(M) = 0$ définissant un plan $\mathcal{P}$. Si F est une fonction définie sur une partie de $\mathbb{R}^2$, l'équation

$$F\left(\left\|\overrightarrow{\Omega M}\right\|^2, P\,(M)\right) = 0$$

définit une surface de révolution (éventuellement vide) dont l'axe est la droite passant par Ω et orthogonale au plan $\mathcal{P}$. Dans un repère orthonormé, $\left\|\overrightarrow{\Omega M}\right\|^2$ est évidemment une expression de la forme

$$\left\|\overrightarrow{\Omega M}\right\|^2 = (x - \alpha)^2 + (y - \beta)^2 + (z - \gamma)^2$$

EXERCICE 21-1.25 Dans l'espace rapporté à un repère orthonormé, équation du cône C de révolution, de sommet $S\,(1, 1, 1)$, d'axe dirigé par $\vec{u}\,(1, 0, 1)$ et passant par $A\,(0, 1, 0)$.

Indication : si $\Delta = S + \mathbf{R}\overrightarrow{u}$ est l'axe du cône, un point $M \neq S$ appartient à $\mathcal{C}$ si et seulement si

$$\frac{d^2(M,\Delta)}{\left\|\overrightarrow{SM}\right\|^2} = \frac{d^2(A,\Delta)}{\left\|\overrightarrow{SA}\right\|^2} = \sin^2\alpha$$

où α est le demi-angle au sommet du cône. On a

$$d^2(M,\Delta) = \frac{\left\|\overrightarrow{u} \wedge \overrightarrow{SM}\right\|^2}{\|\overrightarrow{u}\|^2}$$

puisque $\overrightarrow{u} \wedge \overrightarrow{SM} = \overrightarrow{u} \wedge \overrightarrow{HM}$, où H est la projection orthogonale de M sur Δ. On en déduit aisément l'équation de $\mathcal{C}$.

EXEMPLE 21-1.26 Dans $\mathcal{E}_3$ euclidien, on se donne deux droites $\mathcal{D}_1$ et $\mathcal{D}_2$. On veut décrire la surface $\mathcal{S}$ engendrée par la rotation de $\mathcal{D}_2$ autour de $\mathcal{D}_1$.

- Si les deux droites sont coplanaires, le résultat est simple : c'est un cylindre de révolution si les deux droites sont parallèles, et un cône de révolution (voire un plan) si les deux droites sont concourantes.

- Si $\mathcal{D}_1$ et $\mathcal{D}_2$ ne sont pas coplanaires, on choisit un repère orthonormé $\left(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\right)$ de la manière suivante : l'axe $\mathcal{D}_1$ sera choisi comme axe Oz, le point $O = O_1$ étant pied de la perpendiculaire commune à $\mathcal{D}_1$ et $\mathcal{D}_2$. Nous choisirons cette perpendiculaire commune comme axe Ox de sorte que, si d est la distance de la droite $\mathcal{D}_1$ à $\mathcal{D}_2$, le pied de la perpendiculaire commune sur $\mathcal{D}_2$ est $O_2(d,0,0)$.

  On peut ainsi caractériser $\mathcal{D}_2$ comme intersection de deux plans : le plan parallèle à yOz contenant $\mathcal{D}_2$, d'équation $x = d$, et le plan déterminé par Ox et $\mathcal{D}_2$, d'équation $z = my$ (m est le coefficient directeur, dans le repère $\left(O, \overrightarrow{j}, \overrightarrow{k}\right)$ de l'intersection du plan avec yOz) :

  $$\mathcal{D}_2 : \begin{cases} x = d \\ z = my \end{cases}$$

  Un point $M(x,y,z)$ appartient à $\mathcal{S}$ si et seulement s'il existe

  $$M_0(x_0,y_0,z_0) \in \mathcal{D}_2$$

  vérifiant

  $$x^2 + y^2 = x_0^2 + y_0^2 \quad \text{et} \quad z = z_0$$

  ce qui donne la condition nécessaire et suffisante (si on suppose $m \neq 0$)

  $$x^2 + y^2 = \frac{z^2}{m^2} + d^2$$

  Dans le plan xOz, on trouve donc la méridienne d'équation

  $$x^2 - \frac{z^2}{m^2} = d^2$$

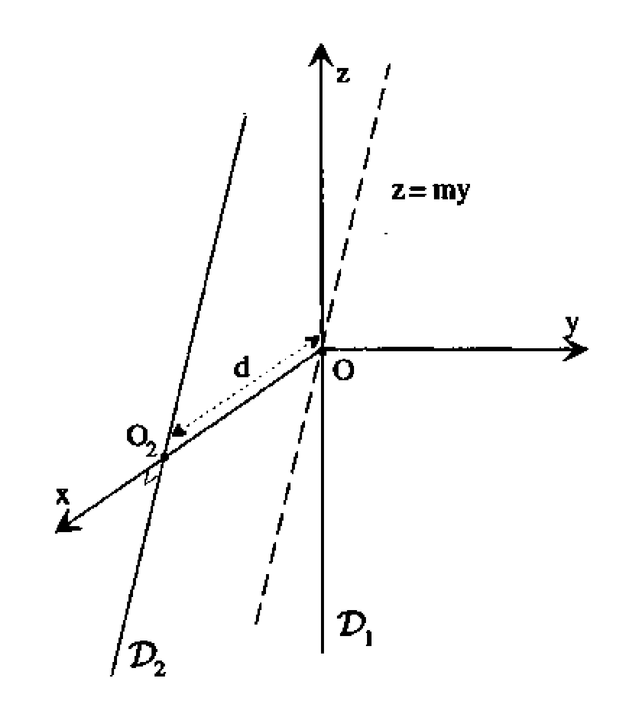


Figure **21.1** – *Choix du repère*

qui est une hyperbole : la surface est un hyperboloïde de révolution (voir section **21-2.2**).

Si $m = 0$, la droite $\mathcal{D}_1$ est orthogonale à $\mathcal{D}_2$, et la condition nécessaire et suffisante est alors simplement $z = 0$ et $x^2 + y^2 \geqslant d^2$: $\mathcal{S}$ est l'ensemble des points de xOy dont la distance à Oz est plus grande que d.

21-2 Surfaces du second degré

21-2.1 Définition

L'étude est analogue à celle de la section **20-1.7.1**

DÉFINITION 21-2.1 *L'espace affine $\mathcal{E}_3$ étant ramené à un repère quelconque $\left(O, \overrightarrow{\imath}, \overrightarrow{\jmath}, \overrightarrow{k}\right)$, on appelle surface du second degré (ou quadrique) tout sous-ensemble $\mathcal{S}$ de $\mathcal{E}_3$ caractérisé par une équation polynomiale de degré* 2

$$\mathcal{S} = \{M(x, y, z) \in \mathcal{E}_3 \mid Q(x, y, z) = 0\}$$

où Q est un polynôme des trois variables x, y et z

$$Q(x, y, z) = A\,x^2 + B\,y^2 + Cz^2 + 2D\,xy + 2E\,yz + 2Fzx + 2Gx + 2Hy + 2Iz + J$$

avec $(A, B, C, D, E, F, G, H, I, J) \in \mathbb{R}^{10}$ et $(A, B, C, D, E, F) \neq 0$

Si Φ est la forme quadratique sur E_3 (espace vectoriel sous-jacent) et l est la forme linéaire dont les matrices dans la base $\left(\vec{I}, \vec{J}, \vec{K}\right)$ sont respectivement

$$M = \begin{pmatrix} A & D & F \\ D & B & E \\ F & E & C \end{pmatrix} \quad \text{et} \quad L = (G, H, I)$$

on a donc

$$M \in \mathcal{S} \Leftrightarrow \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = 0$$

Comme dans le cas de la dimension 2, on vérifie que la forme Φ ne dépend pas du repère choisi. C'est essentiellement en discutant selon la signature de Φ que nous allons obtenir une équation simplifiée (dite réduite) qui va permettre de décrire la surface correspondante.

21-2.2 Classification

On suppose l'espace euclidien, et le repère de départ orthonormé. On cherche un autre repère orthonormé $\left(\Omega, \vec{I}_1, \vec{J}_1, \vec{K}_1\right)$ où l'équation se simplifie :

1. **La signature de Φ est égale à $(3,0)$ ou $(0,3)$** : on dit que $\mathcal{S}$ **est du genre ELLIPSOÏDE**.

   En multipliant l'équation par -1, on peut supposer Φ définie positive. En diagonalisant l'endomorphisme symétrique u associé à Φ (ce qui revient à diagonaliser M), on détermine une base $\left(\vec{I}_1, \vec{J}_1, \vec{K}_1\right)$ avec

   $$\Phi\left(X\vec{I}_1 + Y\vec{J}_1 + Z\vec{K}_1\right) = \mu_1 X^2 + \mu_2 Y^2 + \mu_3 Z^2$$

   (avec tous les $\mu_i > 0$). On peut ensuite, par translation de l'origine, déterminer un repère $\left(\Omega, \vec{I}_1, \vec{J}_1, \vec{K}_1\right)$ où l'équation s'écrira

   $$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \mu_1 X^2 + \mu_2 Y^2 + \mu_3 Z^2 + \delta = 0$$

   Si $\delta > 0$, $\mathcal{S}$ est vide. Elle est réduite à $\{\Omega\}$ si $\delta = 0$. Enfin, si $\delta < 0$, on obtiendra une équation réduite

   $$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1$$

   On dit que $\mathcal{S}$ est un **ellipsoïde de centre** Ω. Dans le cas particulier où u possède une valeur propre double, deux des trois nombres a^2, b^2 et c^2 sont égaux, et l'ellipsoïde possèdera une symétrie de révolution autour d'un des axes de coordonnées.

2. **La signature de Φ est égale à $(2,1)$ ou $(1,2)$** : on dit que $\mathcal{S}$ **est du genre HYPERBOLOÏDE**.

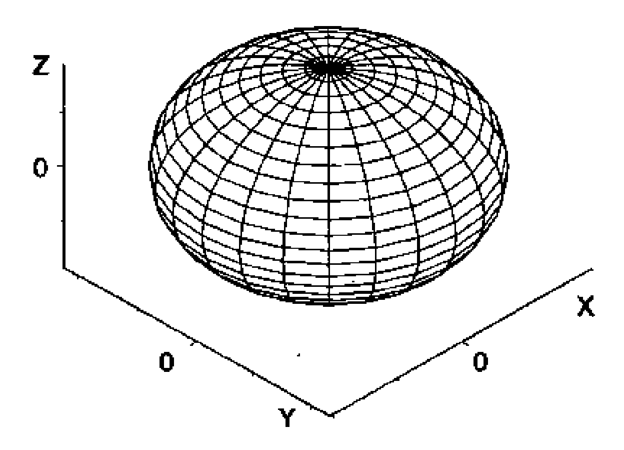


Figure **21.2** – *Ellipsoïde d'équation* $\dfrac{X^2}{a^2}+\dfrac{Y^2}{b^2}+\dfrac{Z^2}{c^2}=1$

On suppose que cette signature vaut (2,1). On aura cette fois, dans une base orthonormale $\left(\overrightarrow{I}_1, \overrightarrow{J}_1, \overrightarrow{K}_1\right)$ diagonalisant u

$$\Phi\left(X\overrightarrow{I}_1+Y\overrightarrow{J}_1+Z\overrightarrow{K}_1\right)=\mu_1X^2+\mu_2Y^2-\mu_3Z^2$$

avec $\mu_i>0$. Par translation de l'origine, on trouvera un repère

$$\left(\Omega, \overrightarrow{I}_1, \overrightarrow{J}_1, \overrightarrow{K}_1\right)$$

où on a

$$M(X,Y,Z)\in\mathcal{S}\Leftrightarrow\mu_1X^2+\mu_2Y^2-\mu_3Z^2+\delta=0$$

La discussion se fait ici selon le signe de δ :

- Si $\delta=0$, on arrive à l'équation réduite

$$M(X,Y,Z)\in\mathcal{S}\Leftrightarrow\frac{X^2}{a^2}+\frac{Y^2}{b^2}=\frac{Z^2}{c^2}$$

qui est l'équation d'un cône de sommet Ω : on dit que $\mathcal{S}$ est un **cône du second degré**. Dans le cas particulier où $\mu_1=\mu_2$ (u possède une valeur propre double), $\mathcal{S}$ possède une symétrie de révolution autour de $\Delta=\Omega+\mathbb{R}\overrightarrow{K}$.

- Si $\delta<0$, en divisant par δ on obtient l'équation réduite

$$M(X,Y,Z)\in\mathcal{S}\Leftrightarrow\frac{X^2}{a^2}+\frac{Y^2}{b^2}=\frac{Z^2}{c^2}+1$$

On dit que $\mathcal{S}$ est un **hyperboloïde à une nappe**. Ici encore, si u possède une valeur propre (strictement positive) double, $\mathcal{S}$ sera une surface de révolution d'axe $\Omega+\mathbb{R}\overrightarrow{K}$.

- Enfin, si $\delta > 0$, $\mathcal{S}$ est un **hyperboloïde à deux nappes** d'équation réduite

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} + \frac{Y^2}{b^2} = \frac{Z^2}{c^2} - 1$$

Si u a une valeur propre double, $\mathcal{S}$ possède une symétrie de révolution.

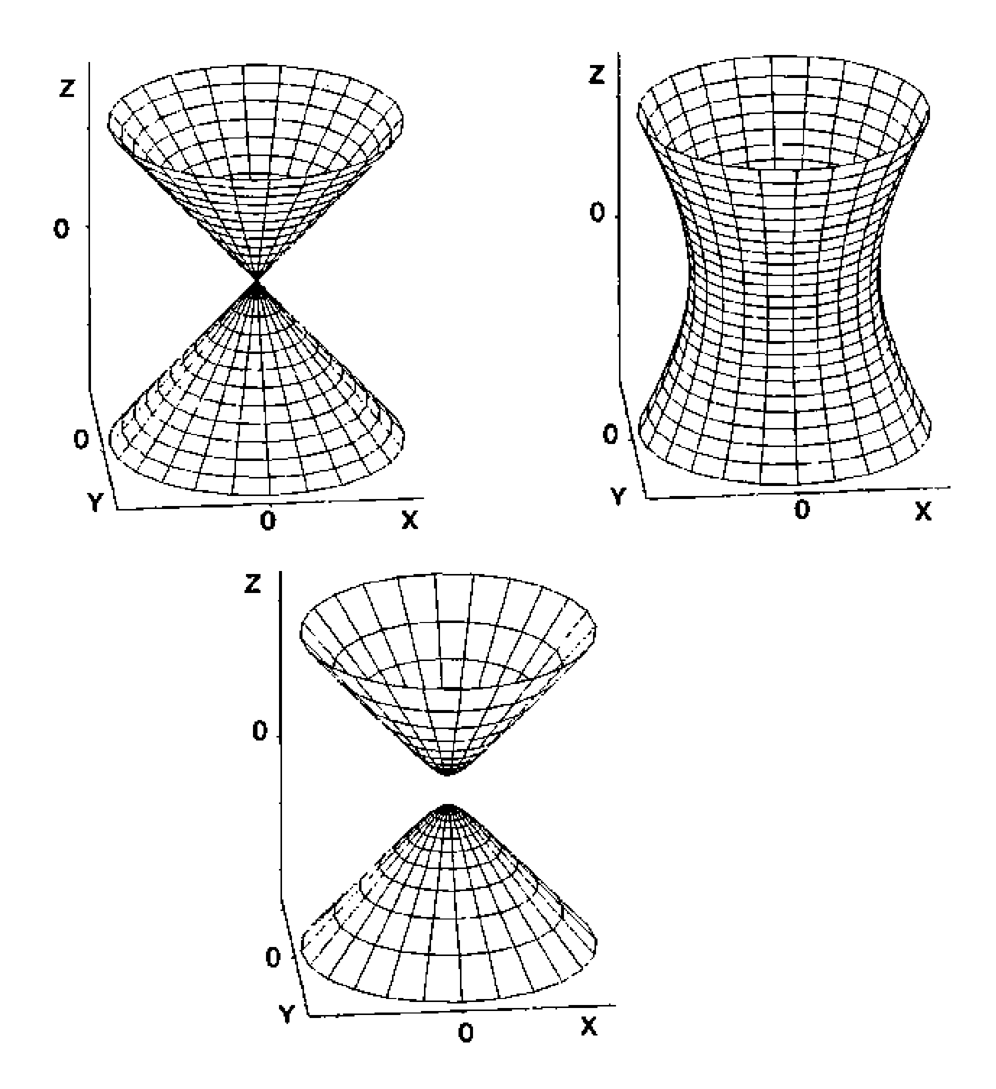
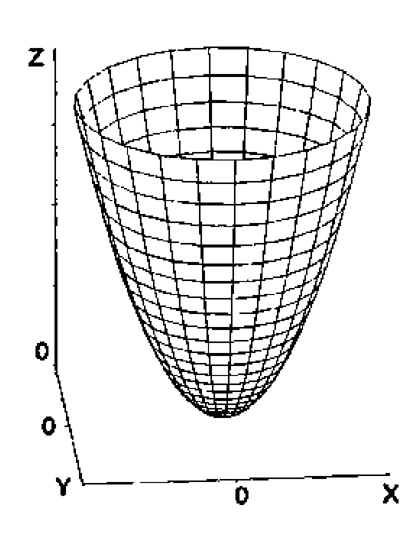


Figure **21.3** – *Signature (2,1) : quadriques du genre hyperboloïde*

3. **La signature de Φ est égale à $(2,0)$ ou $(0,2)$** : on dit que $\mathcal{S}$ **est du genre PARABOLOÏDE ELLIPTIQUE**.

On suppose Φ positive, et on dispose cette fois, en diagonalisant la matrice M, d'une base orthonormale $\left(\overrightarrow{I}_1, \overrightarrow{J}_1, \overrightarrow{K}_1\right)$ avec

$$\Phi\left(X\overrightarrow{I}_1 + Y\overrightarrow{J}_1 + Z\overrightarrow{K}_1\right) = \mu_1 X^2 + \mu_2 Y^2 \quad \text{avec } \mu_1 \text{ et } \mu_2 > 0$$



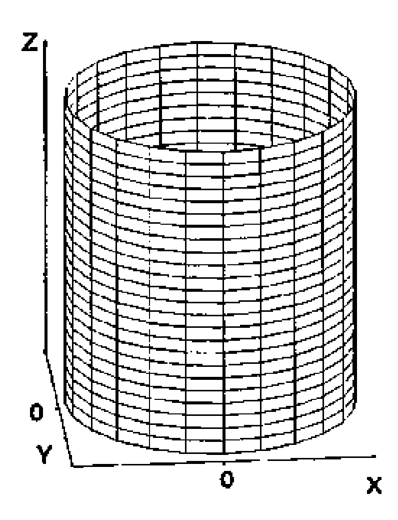


Figure **21.4** – *Signature (2,0) : quadriques du genre paraboloïde elliptique*

Après avoir effectué la rotation des axes du repère, on peut, par translation sur l'origine du repère, faire disparaître les termes de degré 1 en X et Y. On arrive donc à un repère orthonormal $\left(\Omega_1, \overrightarrow{I}_1, \overrightarrow{J}_1, \overrightarrow{K}_1\right)$ où l'on a

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \mu_1 X^2 + \mu_2 Y^2 - 2\alpha Z + \delta = 0$$

- Si $\alpha = 0$, cette équation réduite ne fait pas intervenir la coordonnée Z, et $\mathcal{S}$ est donc un cylindre de génératrices parallèles à $\overrightarrow{K}_1$. Plus précisément, si $\delta > 0$, on aura $\mathcal{S} = \emptyset$, si $\delta = 0$ l'équation de $\mathcal{S}$ équivaut à $X = Y = 0$, et $\mathcal{S}$ est égale à l'axe $\Omega_1 + \mathbb{R}\overrightarrow{K}_1$. Enfin, si $\delta < 0$, on aura un "véritable" cylindre de génératrices parallèles à $\overrightarrow{K}_1$, d'équation réduite

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$$

  Il s'agit donc d'un **cylindre à section droite elliptique**. Il est de révolution si $\mu_1 = \mu_2$.

- Si $\alpha \neq 0$, on peut toujours supposer $\alpha > 0$ (quitte à changer $\overrightarrow{K}_1$ en son opposé). On peut alors translater l'origine en prenant $\Omega = \Omega_1 + \dfrac{\delta}{2\alpha}\overrightarrow{K}_1$. Dans le repère $\left(\Omega, \overrightarrow{I}_1, \overrightarrow{J}_1, \overrightarrow{K}_1\right)$ la surface $\mathcal{S}$ aura une équation réduite de la forme

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} + \frac{Y^2}{b^2} = Z$$

  On dit que $\mathcal{S}$ est un **paraboloïde elliptique.** Si $\mu_1 = \mu_2$, $\Omega + \mathbb{R}\overrightarrow{K}_1$ est un axe de révolution.

4. **La signature de Φ est égale à** $(1,1)$: on dit que $\mathcal{S}$ **est du genre PARABOLOÏDE HYPERBOLIQUE**.

On a cette fois, dans une base orthonormale diagonalisant u

$$\Phi\left(X\vec{I}_1 + Y\vec{J}_1 + Z\vec{K}_1\right) = \mu_1 X^2 - \mu_2 Y^2 \quad \text{avec } \mu_1 \text{ et } \mu_2 > 0$$

Comme dans le cas précédent, on trouve d'abord un repère

$$\left(\Omega_1, \vec{I}_1, \vec{J}_1, \vec{K}_1\right)$$

où on a

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \mu_1 X^2 - \mu_2 Y^2 - 2\alpha Z + \delta = 0$$

- Si $\alpha = 0$, $\mathcal{S}$ est un cylindre de génératrices parallèles à $\vec{K}_1$. Plus précisément, si $\delta = 0$, on aura $M(X,Y,Z) \in \mathcal{S} \Leftrightarrow Y = \pm\sqrt{\frac{\mu_1}{\mu_2}}X$ et $\mathcal{S}$ est réunion de deux plans sécants. Si $\delta \neq 0$, on aura un "véritable" cylindre de génératrices parallèles à $\vec{K}_1$, d'équation réduite (quitte à permuter les rôles joués par X et Y)

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$$

Il s'agit donc d'un **cylindre à section droite hyperbolique**.

- Si $\alpha \neq 0$, on arrivera dans un repère $\left(\Omega, \vec{I}_1, \vec{J}_1, \vec{K}_1\right)$ à l'équation réduite

$$M(X,Y,Z) \in \mathcal{S} \Leftrightarrow \frac{X^2}{a^2} - \frac{Y^2}{b^2} = Z$$

$\mathcal{S}$ est un **paraboloïde hyperbolique**.

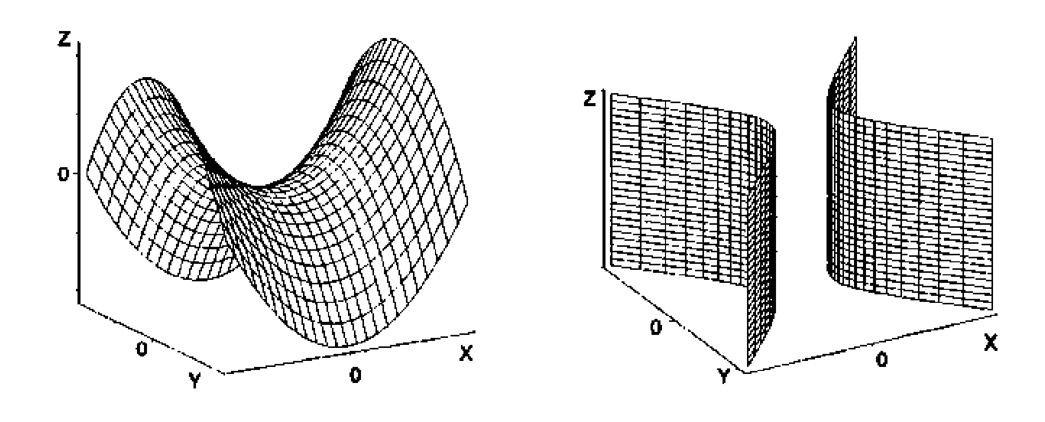


Figure **21.5** – *Signature (1,1) : quadriques du genre paraboloïde hyperbolique*

5. **Enfin, si la signature de Φ est égale à** $(1,0)$ **ou** $(0,1)$: $\mathcal{S}$ **est du genre CYLINDRE PARABOLIQUE.**

   En supposant Φ positive, on a une base orthonormale où

$$\Phi\left(X\overrightarrow{I}_1+Y\overrightarrow{J}_1+Z\overrightarrow{K}_1\right)=\mu_1X^2 \quad \text{avec } \mu_1>0$$

   Dans un repère $\left(\Omega_1,\overrightarrow{I}_1,\overrightarrow{J}_1,\overrightarrow{K}_1\right)$, on arrive à une équation de la forme

$$M(X,Y,Z)\in\mathcal{S}\Leftrightarrow\mu_1X^2-2\alpha Y-2\beta Z+\delta=0$$

   (on est passé de l'origine initiale O à Ω_1 pour faire disparaître le terme du premier degré en X).

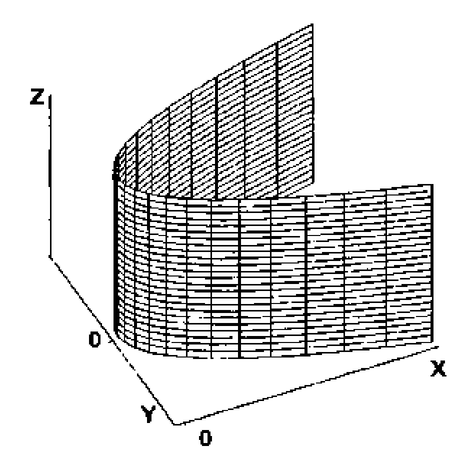


Figure **21.6** – *Signature (1,0) : cylindre parabolique*

- Si $(\alpha,\beta)=(0,0)$, l'équation se réduit à $X^2=-\dfrac{\delta}{\mu_1}$, ce qui donne, en fonction du signe du second membre, $\mathcal{S}=\emptyset$ ou $\mathcal{S}$ est le plan d'équation $X=0$, ou enfin $\mathcal{S}$ est réunion de deux plans parallèles d'équations $X=\pm\sqrt{-\dfrac{\delta}{\mu_1}}$.
- Si $(\alpha,\beta)\neq(0,0)$, on pose

$$\overrightarrow{J_2}=\frac{\alpha\overrightarrow{J}_1+\beta\overrightarrow{K}_1}{\sqrt{\alpha^2+\beta^2}} \quad \text{et} \quad \overrightarrow{K}_2=\frac{-\beta\overrightarrow{J}_1+\alpha\overrightarrow{K}_1}{\sqrt{\alpha^2+\beta^2}}$$

  Dans le repère orthonormé $\left(\Omega_1,\overrightarrow{I}_1,\overrightarrow{J}_2,\overrightarrow{K}_2\right)$, on aura alors

$$M(X,Y,Z)\in\mathcal{S}\Leftrightarrow\mu_1X^2-2\sqrt{\alpha^2+\beta^2}Y+\delta=0$$

  Après translation de l'origine, on arrivera à l'équation réduite

$$X^2=2pY$$

  équation d'un **cylindre à section droite parabolique.**

REMARQUE 21-2.2 Dans la discussion qui précède, nous avons vu qu'une condition suffisante pour qu'une quadrique (non vide) possède une symétrie de révolution est que l'endomorphisme symétrique associé à la forme quadratique ait une valeur propre multiple. On peut démontrer que cette condition est aussi nécessaire (si la quadrique n'est pas réduite à une droite ou un point).

21-2.3 Intersection avec un plan affine

Considérons une quadrique $\mathcal{C}$ de $\mathcal{E}_3$ associée à la forme quadratique $\Phi \neq 0$ et un plan $\mathcal{P}$ déterminé par un point $A \in \mathcal{P}$ et sa direction Π (plan vectoriel de E_3, espace vectoriel sous-jacent). $\mathcal{C}$ possède une équation de la forme

$$M \in \mathcal{C} \Leftrightarrow \Phi\left(\overrightarrow{AM}\right) + 2l_1\left(\overrightarrow{AM}\right) + f = 0$$

avec $f \in \mathbf{R}$ et $l_1 \in E_3^*$. Comme $M \in \mathcal{P} \Leftrightarrow \overrightarrow{AM} \in \Pi$, on en déduit

$$M \in \mathcal{C} \cap \mathcal{P} \Leftrightarrow \overrightarrow{AM} \in \Pi \text{ et } \Phi|_\Pi\left(\overrightarrow{AM}\right) + 2l_1|_\Pi\left(\overrightarrow{AM}\right) + f = 0$$

où $\Phi|_\Pi$ et $l_1|_\Pi$ sont les restrictions respectives de Φ et l_1 au plan Π.

- Si Φ est identiquement nulle[5] en restriction à Π, cette condition se ramène à
$$2l_1|_\Pi\left(\overrightarrow{AM}\right) + f = 0$$
et $\mathcal{C} \cap \mathcal{P}$ peut être $\emptyset$ (si $l_1|_\Pi$ est nulle et $f \neq 0$), le plan $\mathcal{P}$ (si $l_1|_\Pi$ est nulle et $f = 0$, soit $\mathcal{P} \subset \mathcal{C}$) ou une droite de $\mathcal{P}$ (si $l_1|_\Pi$ est non nulle).

- Si $\Phi|_\Pi$ est non nulle, la condition obtenue montre que $\mathcal{C} \cap \mathcal{P}$ est une **courbe du second degré du plan $\mathcal{P}$**, associée à la forme quadratique $\Phi|_\Pi$. Le genre de cette courbe (ellipse, hyperbole ou parabole) dépend de la signature de cette forme quadratique, donc uniquement de la direction du plan $\mathcal{P}$. Il en résulte par exemple que, si $\mathcal{C} \cap \mathcal{P}$ est une ellipse $\mathcal{E}$ et $\mathcal{P}_1$ est un plan parallèle à $\mathcal{P}$, l'intersection $\mathcal{C} \cap \mathcal{P}_1$ peut être une ellipse de même excentricité que celle de $\mathcal{E}$, avec des axes parallèles à ceux de $\mathcal{E}$, ou un point ou l'ensemble vide.

EXERCICE 21-2.3 Dans $\mathcal{E}_3$ rapporté à un repère orthonormé, on considère le cône de révolution $\mathcal{C}$ d'équation

$$z^2 = \tan^2\alpha\,(x^2 + y^2)$$

Soit $\mathcal{P}$ un plan dont la normale fait un angle β avec l'axe Oz. Discuter, en fonction de β, le genre de l'intersection $\mathcal{C} \cap \mathcal{P}$. Dans quels cas cette intersection n'est-elle pas une "vraie" conique ?

[5]Ceci n'est possible que si la signature de Φ est $(1,1)$, $(1,0)$ ou $(0,1)$. Sinon, il existe au moins un plan vectoriel Π_1 sur lequel Φ est définie (positive ou négative), et l'intersection $\Pi \cap \Pi_1$ contient au moins une droite vectorielle.

EXERCICE 21-2.4 $\mathcal{E}_3$ est rapporté à un repère orthonormé. Déterminer les plans affines coupant l'hyperboloïde d'équation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z^2}{c^2}+1$$

(avec $a>b>0$) selon un cercle.

Indication : il s'agit d'abord de déterminer les plans vectoriels sur lesquels la forme Φ induit une forme quadratique proprtionnelle au carré de la norme euclidienne.

21-2.4 Remarques

21-2.4.1 Quadriques réglées

On dit qu'une surface est réglée si elle est réunion d'une famille de droites. Pour les quadriques, parmi les cas discutés précédemment, les cônes du second degré, les cylindres (à section droite elliptique, hyperbolique ou parabolique), les plans, les couples de plans (parallèles ou sécants) et le cas tout à fait particulier d'une seule droite sont des quadriques réglées.

Il reste à étudier les ellipsoïdes, hyperboloïdes (à une ou deux nappes) et paraboloïdes (elliptiques ou hyperboliques) :

- Un ellipsoïde est clairement compact, et ne contient donc aucune droite. On peut même être plus précis en disant qu'il ne contient aucun segment de droite (non réduit à un point). Cette propriété, presque une évidence, peut se justifier en disant que l'ellipsoïde d'équation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

est image de la sphère de centre O et de rayon 1 par la bijection affine

$$M(x,y,z)\mapsto M'(ax,by,cz)$$

et en remarquant que la sphère unité d'un espace euclidien est toujours **strictement convexe :** si $\|\overrightarrow{u}\|=\|\overrightarrow{v}\|=1$, on ne peut avoir

$$\left\|\frac{\overrightarrow{u}+\overrightarrow{v}}{2}\right\|=1$$

si $\overrightarrow{u}\neq\overrightarrow{v}$ (exercice).

- Un paraboloïde elliptique $\mathcal{P}_e$ d'équation réduite

$$z=\frac{x^2}{a^2}+\frac{y^2}{b^2}$$

ne contient aucune droite : une telle droite ne pourrait se trouver dans un plan où z est constant, puisque l'intersection de $\mathcal{P}_e$ avec un tel plan est une ellipse (un point ou $\emptyset$). Comme $\mathcal{P}_e$ est inclus dans le demi-espace $z\geqslant 0$, il ne peut également contenir une droite non parallèle à xOy. On peut dire de plus qu'aucun segment non réduit à un point n'est inclus dans $\mathcal{P}_e$.

Ceci peut se prouver en vérifiant que tout point de $\mathcal{P}_e$ est elliptique : si $f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$, la forme quadratique associée à la matrice Hessienne de f en tout point est égale à $2f$ et est définie positive. La surface est donc située strictement au dessus de son plan tangent[6] en tout point, alors qu'un segment tracé sur $\mathcal{P}_e$ serait contenu dans le plan tangent en chacun de ses points.

- Un hyperboloïde $\mathcal{H}_2$ à deux nappes, d'équation réduite

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} - 1$$

ne contient de même aucune droite : une telle droite serait nécessairement incluse dans un des demi-espaces $z \geqslant c$ ou $z \leqslant c$, et serait donc parallèle au plan xOy. Comme précédemment, ce n'est pas possible. On pourrait également montrer que tous les points de $\mathcal{H}_2$ sont elliptiques, en étudiant la matrice Hessienne de la fontion

$$f(x,y) = \pm c\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1}$$

Il n'y a donc aucun segment de droite inclus dans $\mathcal{H}_2$.

- Le paraboloïde hyperbolique $\mathcal{P}_h$ d'équation réduite

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$

est réglé :

Cherchons d'abord les droites horizontales incluses dans $\mathcal{P}_h$: l'intersection avec un plan horizontal est une hyperbole (dégénérée en deux droites si ce plan est xOy). Il y a donc exactement deux droites horizontales dans $\mathcal{P}_h$. Elles ont pour équations

$$\mathcal{D}_\infty : \begin{cases} \frac{x}{a} + \frac{y}{b} = 0 \\ z = 0 \end{cases} \quad \text{et} \quad \mathcal{D}'_\infty : \begin{cases} \frac{x}{a} - \frac{y}{b} = 0 \\ z = 0 \end{cases}$$

Si une droite $\mathcal{D}$ n'est pas parallèle à xOy, elle peut être paramétrée par z sous la forme

$$\begin{cases} x = \alpha z + \beta \\ y = \alpha' z + \beta' \end{cases}$$

Les quantités $\left(\frac{x}{a} - \frac{y}{b}\right)$ et $\left(\frac{x}{a} + \frac{y}{b}\right)$ sont donc fonctions affines de z. Si $\mathcal{D} \subset \mathcal{P}_h$ la condition

$$z = \left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right)$$

[6]Localement, mais aussi globalement, puisque f étant un polynôme de degré 2, le développement limité de f à l'ordre 2 en tout point ne contient pas de terme complémentaire.

(vérifiée pour tout z) montre qu'une de ces quantités est constante non nulle, l'autre étant proportionnelle à z. On trouve donc, pour $\lambda \in \mathbf{R}^*$ les droites $\mathcal{D}_\lambda$ et $\mathcal{D}'_\lambda$ d'équations respectives

$$\mathcal{D}_\lambda : \begin{cases} \dfrac{x}{a} - \dfrac{y}{b} = \lambda z \\[2ex] \dfrac{x}{a} + \dfrac{y}{b} = \dfrac{1}{\lambda} \end{cases} \quad \text{et} \quad \mathcal{D}'_\lambda : \begin{cases} \dfrac{x}{a} + \dfrac{y}{b} = \lambda z \\[2ex] \dfrac{x}{a} - \dfrac{y}{b} = \dfrac{1}{\lambda} \end{cases}$$

incluses dans $\mathcal{P}_h$. On peut considérer que les droites $\mathcal{D}_\infty$ et $\mathcal{D}'_\infty$ sont les positions limites de ces droites pour $\lambda \to +\infty$. On remarquera que toutes les droites $\mathcal{D}_\lambda$ sont parallèles au plan d'équation $\dfrac{x}{a} + \dfrac{y}{b}$, avec une propriété analogue pour les droites $\mathcal{D}'_\lambda$.

Par chaque point $M\,(x, y, z)$ de $\mathcal{P}_h$ passe une unique droite de la forme $\mathcal{D}_\lambda$ et une unique droite de la forme $\mathcal{D}'_\mu$. L'intersection de la surface avec le plan tangent en M est exactement la réunion $\mathcal{D}_\lambda \cap \mathcal{D}'_\mu$. On vérifie facilement que chaque point de $\mathcal{P}_h$ est hyperbolique.

- On pourrait faire une étude analogue pour l'hyperboloïde à une nappe $\mathcal{H}_1$ d'équation réduite

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} + 1$$

en écrivant cette équation sous la forme

$$\left(\frac{x}{a} + \frac{z}{c}\right)\left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 - \frac{y}{b}\right)\left(1 + \frac{y}{b}\right)$$

On trouverait deux familles de droites

$$\mathcal{D}_\lambda : \begin{cases} \dfrac{x}{a} + \dfrac{z}{c} = \lambda\left(1 - \dfrac{y}{b}\right) \\[2ex] 1 + \dfrac{y}{b} = \lambda\left(\dfrac{x}{a} - \dfrac{z}{c}\right) \end{cases} \quad \text{et} \quad \mathcal{D}'_\lambda : \begin{cases} \dfrac{x}{a} + \dfrac{z}{c} = \lambda\left(1 + \dfrac{y}{b}\right) \\[2ex] 1 - \dfrac{y}{b} = \lambda\left(\dfrac{x}{a} - \dfrac{z}{c}\right) \end{cases}$$

complétées par

$$\mathcal{D}_\infty : \begin{cases} 1 - \dfrac{y}{b} = 0 \\[2ex] \dfrac{x}{a} - \dfrac{z}{c} = 0 \end{cases} \quad \text{et} \quad \mathcal{D}'_\infty : \begin{cases} 1 + \dfrac{y}{b} = 0 \\[2ex] \dfrac{x}{a} - \dfrac{z}{c} = 0 \end{cases}$$

On vérifierait que, par chaque point de $\mathcal{H}_1$, passe une et une seule droite de chaque famille.

On peut aussi remarquer qu'un hyperboloïde de révolution à une nappe est engendré par la rotation d'une droite autour de son axe : nous avons vu à l'exemple **21-1**.26 que la surface d'équation (en repère orthonormé)

$$x^2 + y^2 = \frac{z^2}{m^2} + d^2$$

était engendrée par la rotation (autour de Oz) de la droite d'équation

$$\mathcal{D}: \begin{cases} x = d \\ z = my \end{cases}$$

(mais aussi par la droite $\mathcal{D}'$ obtenue en changeant m en son opposé). En tournant, $\mathcal{D}$ et $\mathcal{D}'$ décrivent les deux familles de génératrices. Comme $\mathcal{H}_1$ se déduit d'un hyperboloïde de révolution par affinité, on retrouve le fait que $\mathcal{H}_1$ est une surface réglée.

21-2.4.2 Image par une bijection affine

Considérons une quadrique $\mathcal{C}$ d'équation

$$\Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = 0$$

et considérons une bijection affine

$$f : \mathcal{E}_3 \to \mathcal{E}_3 \quad M \mapsto M' = f(M)$$

$u \in \mathcal{GL}(E_3)$ étant l'application linéaire associée : si O' est l'image de O, on a

$$\forall M \in \mathcal{E}_3 \quad \overrightarrow{O'M'} = u\left(\overrightarrow{OM}\right)$$

Nous cherchons à déterminer la nature de l'ensemble $f(\mathcal{C})$.

On a

$$M' \in f(\mathcal{C}) \Leftrightarrow \exists M \in \mathcal{C} \quad \overrightarrow{OM} = u^{-1}\left(\overrightarrow{O'M'}\right)$$

On obtient donc

$$M' \in f(\mathcal{C}) \Leftrightarrow \Phi \circ u^{-1}\left(\overrightarrow{O'M'}\right) + 2l \circ u^{-1}\left(\overrightarrow{O'M'}\right) + J = 0$$

On en déduit que $f(\mathcal{C})$ est une quadrique qui est associée à la forme quadratique $\Psi = \Phi \circ u^{-1}$. Comme la signature d'une forme quadratique est définie en considérant les sous-espaces sur lesquels la forme est définie positive ou négative, il est clair que $\Phi \circ u^{-1}$ a même signature que Φ et $f(\mathcal{C})$ est donc de même genre que $\mathcal{C}$. De plus, f étant une bijection affine transforme toute droite de $\mathcal{E}_3$ en une droite. Cette remarque, et une étude attentive de la discussion menée lors de la classification, permet de voir que $\mathcal{C}$ et $f(\mathcal{C})$ sont exactement de même nature :

Par exemple, si la signature de Φ est $(2,1)$, $\mathcal{C}$ peut être un cône du second degré ou un hyperboloïde à une ou deux nappes. Des propriétés, conservées par transformation affine, permettent de distinguer ces différents cas :

- si $\mathcal{C}$ est un hyperboloïde à deux nappes, il est non connexe, et il en est de même de $f(\mathcal{C})$ (car f est bicontinue).
- Si $\mathcal{C}$ est réglée, il en est de même de $f(\mathcal{C})$. On distingue le cône de l'hyperboloïde à une nappe en remarquant que, dans le premier cas, les génératrices passent par un point fixe et donc $f(\mathcal{C})$ est aussi un cône.

PROPOSITION 21-2.5 Si $\mathcal{C}$ est une quadrique de $\mathcal{E}_3$ et f est une bijection affine de $\mathcal{E}_3$ dans lui-même, $f(\mathcal{C})$ est une quadrique de même nature que $\mathcal{C}$.

Cette proposition va nous permettre de déterminer la nature exacte d'une quadrique sans calculer l'équation réduite. La méthode est décrite sur un exemple :

EXEMPLE 21-2.6 Nature de la quadrique $\mathcal{C}$, dont l'équation dans un repère orthonormé $\mathcal{R}$ de $\mathcal{E}_3$ est

$$x^2 + 3y^2 - 8z^2 - 4xy + 2xz - 10yz - x - 5y - 12z + 1 = 0$$

La méthode de Gauss permet d'écrire la forme quadratique associée sous la forme

$$\Phi(x,y,z) = x^2 + 3y^2 - 8z^2 - 4xy + 2xz - 10yz = (x - 2y + z)^2 - (y + 3z)^2$$

Elle est donc de signature $(1,1)$, et $\mathcal{C}$ peut donc être un paraboloïde hyperbolique, un cylindre à section droite hyperbolique ou la réunion de deux plans sécants. Si on complète le système de deux formes linéaires intervenant dans la décomposition de Φ par une troisième forme affine indépendante, en posant[7]

$$(*)\begin{cases} X = x - 2y + z \\ Y = y + 3z \\ Z = x + 5y + 12z - 1 \end{cases}$$

l'équation s'écrit alors

$$X^2 - Y^2 = Z$$

Dans un repère orthonormé, ce serait l'équation **réduite** d'un paraboloïde hyperbolique. Ici, on peut interpréter les formules $(*)$ comme représentant un changement de repère. Mais comme la matrice associée n'est pas orthogonale, le nouveau repère n'est pas orthonormé.

Pour voir que $\mathcal{C}$ est effectivement un paraboloïde hyperbolique, on interprète plutôt $(*)$ comme définissant une bijection affine

$$g : \mathcal{E}_3 \to \mathcal{E}_3 \quad M(x,y,z) \mapsto P(X,Y,Z)$$

(les coordonnées étant exprimées dans le repère de départ $\mathcal{R}$). Si $\mathcal{P}_h$ est le paraboloïde hyperbolique dont l'équation dans $\mathcal{R}$ est

$$x^2 - y^2 = z$$

on a $\mathcal{C} = g^{-1}(\mathcal{P}_h)$, et $\mathcal{C}$ est bien un paraboloïde hyperbolique.

REMARQUE 21-2.7 Ce qui précède peut aussi être utilisé dans le plan pour déterminer rapidement la nature exacte d'une courbe du second degré.

---

[7] La matrice

$$M = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 3 \\ 1 & 5 & 12 \end{pmatrix}$$

est en effet inversible. Si la troisième ligne se trouvait être combinaison linéaire des deux premières, l'équation de $\mathcal{C}$ serait dégénérée (ne ferait intervenir que deux coordonnées dans un bon repère) et $\mathcal{C}$ serait un cylindre.

21-2.4.3 Centre de symétrie

PROPOSITION 21-2.8 Soit $\mathcal{C}$ une quadrique non vide d'équation

$$\Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = 0$$

Le point $\Omega \in \mathcal{E}_3$ est centre de symétrie de $\mathcal{C}$ si et seulement si

$$\exists K \in \mathbb{R} \quad \forall M \in \mathcal{E}_3 \quad \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = \Phi\left(\overrightarrow{\Omega M}\right) + K$$

Démonstration : Cette condition est évidemment suffisante pour que Ω soit un centre de symétrie, car il est alors évident que

$$M \in \mathcal{C} \Leftrightarrow M' = \Omega + \overrightarrow{M\Omega} \in \mathcal{C}$$

Supposons réciproquement que Ω soit un centre de symétrie. Nous savons qu'il existe une forme linéaire l_1 et un réel K tels que

$$\forall M \in \mathcal{E}_3 \quad \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = \Phi\left(\overrightarrow{\Omega M}\right) + 2l_1\left(\overrightarrow{\Omega M}\right) + K$$

Comme Ω est centre de symétrie, si le point M vérifie l'équation de $\mathcal{C}$, il en est de même du point M' tel que $\overrightarrow{\Omega M'} = -\overrightarrow{\Omega M}$. On en déduit que

$$\forall M \in \mathcal{C} \quad \Phi\left(\overrightarrow{\Omega M}\right) + 2l_1\left(\overrightarrow{\Omega M}\right) + K$$
$$= \Phi\left(\overrightarrow{\Omega M}\right) - 2l_1\left(\overrightarrow{\Omega M}\right) + K = 0$$

ce qui montre que $M \in \mathcal{C} \Rightarrow l_1\left(\overrightarrow{\Omega M}\right) = 0$ et donc

$$\mathcal{C} \subset \Omega + \ker l_1$$

Si la forme l_1 n'est pas nulle, cela signifie que $\mathcal{C}$ est incluse dans un plan affine. En reprenant l'étude menée lors de la classification, ceci nous montre que $\mathcal{C}$ est un point, une droite ou un plan. Le point Ω appartient donc à $\mathcal{C}$ et, dans ces trois cas, nous avons vu dans la discussion de la section **21-2.2** qu'en ramenant l'origine en Ω on faisait disparaître les termes du premier degré et le terme constant, c'est à dire

$$\forall M \in \mathcal{E}_3 \quad \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J = \Phi\left(\overrightarrow{\Omega M}\right)$$

La forme l_1 est donc toujours nulle, et on a alors le résultat annoncé.
■

Pour déterminer une équation réduite on cherche souvent, dans la pratique, à ramener d'abord l'origine du repère en un éventuel centre de symétrie, avant d'effectuer une rotation des axes. Comme

$$\forall M \in \mathcal{E}_2 \quad \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J$$
$$= \Phi\left(\overrightarrow{\Omega M}\right) + 2\left[l\left(\overrightarrow{\Omega M}\right) + \varphi\left(\overrightarrow{O\Omega}, \overrightarrow{\Omega M}\right)\right] + K + \Phi\left(\overrightarrow{O\Omega}\right) + 2l\left(\overrightarrow{O\Omega}\right)$$

(φ étant la forme polaire de Φ) le point Ω est donc un centre de symétrie si et seulement si

$$\varphi\left(\overrightarrow{O\Omega},\bullet\right) = -l$$

- Si Φ est non dégénérée (genre ellipsoïde ou hyperboloïde), l'application $\overrightarrow{x} \mapsto \varphi(\overrightarrow{x},\bullet)$ est un isomorphisme de E_3 vers son dual, et il existe un unique Ω vérifiant cette équation. La quadrique a un unique centre de symétrie.
- Si Φ est de rang 2, le radical de φ est de dimension 1, et lorsque l'équation

  $$\varphi(\overrightarrow{x},\bullet) = -l$$

  possède des solutions, celles-ci forment une droite vectorielle. Il en résulte que la quadrique peut ne pas avoir de centre de symétrie (paraboloïde elliptique ou hyperbolique), ou posséder une droite affine de centres de symétrie (cylindres à section droite elliptique ou hyperbolique, droite ou réunion de deux plans sécants).
- Enfin, si Φ est de rang 1, il n'y a pas de centre de symétrie (cylindre parabolique) ou un plan formé de centres de symétrie ($\mathcal{C}$ est un plan ou un couple de plans parallèles).

Dans la pratique :

Si $F : \mathcal{E}_3 \to \mathbb{R}$ est définie par $F(M) = \Phi\left(\overrightarrow{OM}\right) + 2l\left(\overrightarrow{OM}\right) + J$, on a clairement

$$dF_M = 2\left(\varphi\left(\overrightarrow{OM},\bullet\right) + l\right)$$

et donc les points critiques de F sont exactement les centres de symétrie de la quadrique :

PROPOSITION 21-2.9 Si une quadrique non vide $\mathcal{C}$ a, dans un repère quelconque, une équation du second degré

$$Q(x,y,z) = 0$$

un point $\Omega(x_0,y_0,z_0)$ est un centre de symétrie de $\mathcal{C}$ si et seulement si

$$\frac{\partial Q}{\partial x}(x_0,y_0,z_0) = \frac{\partial Q}{\partial y}(x_0,y_0,z_0) = \frac{\partial Q}{\partial z}(x_0,y_0,z_0) = 0$$

21-3 Exercices

Sauf mention contraire, les équations des surfaces intervenant dans les exercices qui suivent sont écrites dans un repère orthonormé.

EXERCICE 21-3.1 Soit E un espace vectoriel euclidien de dimension n et q une forme quadratique sur E, u l'endomorphisme symétrique associé à q. Montrer qu'il existe une base orthonormale de E formée de vecteurs annulant q ssi la trace de u est nulle. Montrer qu'alors il y a une infinité de telles bases.
Soit Γ la conique de $\mathbb{R}^3$ euclidien d'équations $ax^2 + by^2 - 1 = 0$ et $z = 0$. Déterminer le lieu des points M tels qu'il existe un trièdre rectangle d'origine M s'appuyant sur Γ.

EXERCICE 21-3.2 Nature de la quadrique d'équation

$$(b^2+c^2)x^2+(c^2+a^2)y^2+(a^2+b^2)z^2-2bcyz-2cazx-2abxy-d=0$$

EXERCICE 21-3.3 Soient $a>b>0$. Déterminer les plans coupant

$$(H):\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

suivant des cercles. Quel est le lieu des centres de ces cercles ?

EXERCICE 21-3.4 Déterminer l'ensemble des foyers des paraboles tracées sur un cône de révolution.

EXERCICE 21-3.5 Déterminer l'équation du cylindre de génératrices parallèles au vecteur $\vec{a}(u,v,w)$ circonscrit à la surface d'équation

$$(E):\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

Ce cylindre peut-il être de révolution ? Quel est alors son rayon ?

EXERCICE 21-3.6 Soit (S) la sphère d'équation

$$x^2+y^2+z^2-2cz-R^2=0$$

et Q_λ la quadrique d'équation

$$x^2+y^2+(1+\lambda)z^2-2cz-R^2=0$$

Un plan tangent (P) à la sphère rencontre Q_λ suivant une conique Γ. Montrer que $F=(P)\cap(S)$ est un foyer de Γ.

EXERCICE 21-3.7 Déterminer le lieu des points équidistants d'une droite et d'un plan. Déterminer le lieu des points équidistants de deux droites données. Si ces droites ont pour équations $(x=0,y=0)$ et $(z=0,x+y=1)$ déterminer les droites incluses dans ce lieu. Si D et D' sont deux droites non coplanaires, un point M se projette orthogonalement en H sur D et en H' sur D'. Trouver le lieu de M pour que

$$(MH)^2+(M'H')^2=a^2$$

EXERCICE 21-3.8 Soit Σ la surface d'équation cartésienne $(x^2+y^2)z^2=a^2y^2$. Quelle est la nature de l'intersection de Σ avec un plan parallèle à xOy ? Quelles sont les droites incluses dans Σ ? Montrer que ces droites restent tangentes à deux sphères fixes. Déterminer les trajectoires orthogonales de ces droites (c'est à dire les arcs tracés sur la surface coupant toutes ces droites selon un angle droit).

EXERCICE 21-3.9 Soient a,b,c trois réels tels que $a>b>c$. Pour tout réel λ, on définit Σ_λ ensemble des points de coordonnées (x,y,z) tels que

$$\frac{x^2}{a+\lambda}+\frac{y^2}{b+\lambda}+\frac{z^2}{c+\lambda}=1$$

1. Pour quels λ Σ_λ est-il non vide ?
2. Pour quels λ et μ tels que $\lambda<\mu$, l'intersection $\Sigma_\lambda\cap\Sigma_\mu$ est-elle non vide ?
3. Montrer qu'alors, en tout point de l'intersection, les plans tangents aux deux surfaces sont perpendiculaires.

Chapitre 22

Aide-mémoire de géométrie analytique

Le plan $\mathcal{E}_2$ et l'espace $\mathcal{E}_3$ sont euclidiens orientés. Les repères $(O, \vec{\imath}, \vec{\jmath})$ et $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ sont orthonormés directs.

- **1 <u>Equation normale d'une droite, d'un plan</u>**

Pour une droite $\mathcal{D}$ du plan, c'est une équation de la forme

$$x\cos\theta + y\sin\theta - p = 0$$

Le vecteur unitaire

$$\vec{u} = \cos\theta\,\vec{\imath} + \sin\theta\,\vec{\jmath}$$

est alors normal à $\mathcal{D}$. Si H est la projection orthogonale de O sur $\mathcal{D}$, on a $\overrightarrow{OH} = p\vec{u}$. Si $M(x,y)$ est un point du plan, la quantité $x\cos\theta + y\sin\theta - p$ représente donc

$$\overrightarrow{OM}.\vec{u} - \overrightarrow{OH}.\vec{u} = \overrightarrow{HM}.\vec{u}$$

C'est donc la **"distance" algébrique de M à $\mathcal{D}$.**
On a le même résultat dans l'espace pour un plan $\mathcal{P}$ d'équation

$$\alpha x + \beta y + \gamma z - p = 0$$

si $\alpha^2 + \beta^2 + \gamma^2 = 1$ Le vecteur $\vec{n} = \alpha\vec{\imath} + \beta\vec{\jmath} + \gamma\vec{k}$ est unitaire et oriente la normale au plan.

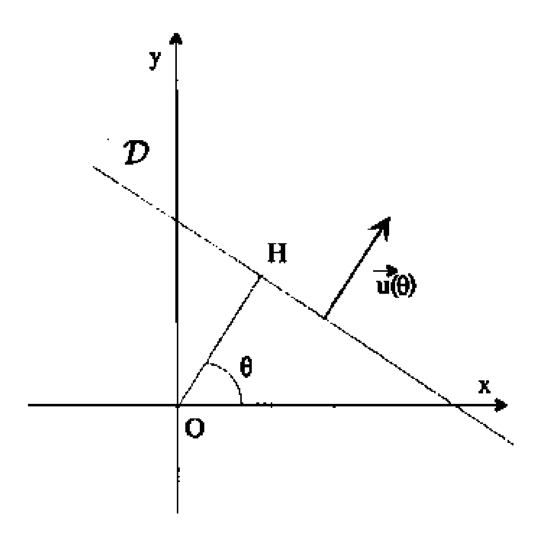


Figure **22.1** – *Vecteur unitaire normal à une droite*

- **2 Distance algébrique d'un point à un plan**

Conséquence de ce qui précède : si $ux + vy + wz + h = 0$ est l'équation du plan $\mathcal{P}$ et $M(x, y, z)$ a pour projection orthogonale m sur $\mathcal{P}$ on a :

$$\overline{mM} = \overrightarrow{mM}.\vec{n} = \frac{ux + vy + wz + h}{\sqrt{u^2 + v^2 + w^2}}$$

si $\vec{n} = \dfrac{\vec{N}}{\|\vec{N}\|}$ avec $\vec{N} = u\vec{\imath} + v\vec{\jmath} + w\vec{k}$ oriente la normale. La **direction du plan** est évidemment le plan vectoriel d'équation $ux + vy + wz = 0$.

- **3 Condition pour que 3(4) points soient alignés (coplanaires)**

Dans le plan les points $(M_i(x_i, y_i))_{i=1...3}$ sont alignés ssi

$$\det(\overrightarrow{M_1M_2}, \overrightarrow{M_1M_3}) = 0$$

Cette condition peut s'écrire de manière plus symétrique sous la forme

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

Condition analogue dans l'espace pour que 4 points soient coplanaires. L'équation du plan passant par trois points A, B et C s'écrira notamment

$$\begin{vmatrix} x & x_A & x_B & x_C \\ y & y_A & y_B & y_C \\ z & z_A & z_B & z_C \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0$$

Dans le cas particulier où $A(a,0,0), B(0,b,0)$ et $C(0,0,c)$ sont les intersections du plan avec les axes de coordonnées, on peut écrire directement cette équation

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$$

- 4 **Faisceau linéaire de plans (de droites)**

Dans l'espace on considère **deux plans distincts** $\mathcal{P}_1$ et $\mathcal{P}_2$ d'équations affines $P_1(M)=0$ et $P_2(M)=0$. ($\mathcal{P}_1 \neq \mathcal{P}_2 \Longrightarrow$ les formes P_1 et P_2 ne sont pas proportionnelles.)

- Si **les deux plans sont parallèles**, tout plan parallèle à leur direction commune a une équation de la forme $\lambda_1 P_1(M)+\lambda_2 P_2(M)=0$ où $(0,0)\neq(\lambda_1,\lambda_2)\in\mathbb{R}^2$. Plus précisément si M_0 est un point de l'espace, le plan d'équation

$$P_2(M_0)P_1(M)-P_1(M_0)P_2(M)=0$$

est parallèle à la direction commune de $\mathcal{P}_1$ et $\mathcal{P}_1$ et passe par M_0.

- Si **les deux plans sont sécants suivant une droite** $\mathcal{D}$, tout plan contenant $\mathcal{D}$ a une équation de la forme $\lambda_1 P_1(M)+\lambda_2 P_2(M)=0$. Si $M_0 \notin \mathcal{D}$, le plan $(M_0,\mathcal{D})$ a pour équation

$$P_2(M_0)P_1(M)-P_1(M_0)P_2(M)=0$$

Il s'agit ici d'une application directe d'un résultat de dualité : si φ_1 et φ_2 sont deux formes linéaires sur un espace vectoriel, une forme linéaire φ est combinaison linéaire de φ_1 et φ_2 ssi

$$\mathrm{Ker}(\varphi_1)\cap\mathrm{Ker}(\varphi_2)\subset\mathrm{Ker}(\varphi)$$

Il suffit alors d'écrire les équations des plans, en choisissant un point $M_0\in\mathcal{P}_1\cap\mathcal{P}_2$, sous la forme

$$P_1(M)=\varphi_1(\overrightarrow{M_0M})=0 \text{ et } P_2(M)=\varphi_1(\overrightarrow{M_0M})=0$$

L'intersection des deux plans a pour direction l'intersection des noyaux des formes φ_1 et φ_2. Un plan contenant cette intersection aura une équation de la forme

$$P(M)=\varphi(\overrightarrow{M_0M})=0$$

avec φ combinaison linéaire de φ_1 et φ_2.

On a évidemment le même résultat dans le plan avec deux droites distinctes parallèles ou concourantes en un point. Par exemple, si $P_1(M)=0$ et $P_2(M)=0$ sont les équations normales de deux droites concourantes, les bissectrices ont pour équation

$$P_1(M)\pm P_2(M)=0$$

ce qui peut se justifier par le fait que les bissectrices sont le lieu des points équidistants des deux droites, ou parce que, si $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$ sont unitaires normaux aux deux droites, $\overrightarrow{u_1} \pm \overrightarrow{u_2}$ sont deux vecteurs orthogonaux qui dirigent les deux bissectrices.

- **5 Distance d'un point à une droite de l'espace**

  – Si la droite $\mathcal{D}$ est définie par **un point** A **et un vecteur directeur** $\vec{u}$, on a

$$d(M,\mathcal{D}) = \frac{\|\overrightarrow{AM} \wedge \vec{u}\|}{\|\vec{u}\|}$$

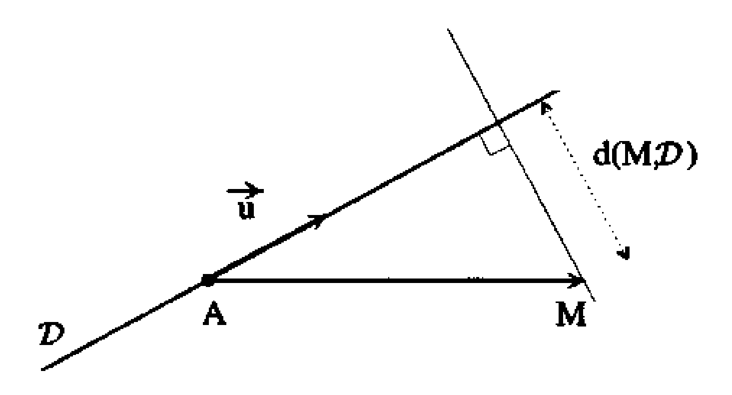


Figure **22.2** – *Distance d'un point à une droite*

  – Si la droite est définie comme **intersection de deux plans d'équations** $P_1(M) = 0$ **et** $P_2(M) = 0$, on a immédiatement des vecteurs normaux aux plans $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$:

    * Si $\overrightarrow{u_1}.\overrightarrow{u_2}=0$, les deux plans sont perpendiculaires et le théorème de Pythagore donne immédiatement

$$\forall M \in \mathcal{E}_3 \quad d^2(M,\mathcal{D}) = d^2(M,\mathcal{P}_1) + d^2(M,\mathcal{P}_2)$$

    * Sinon on détermine α avec $\overrightarrow{u_1}.(\overrightarrow{u_1} + \alpha\overrightarrow{u_2}) = 0$ et on remplace le plan $\mathcal{P}_2$ par le plan $\mathcal{P}_2'$ d'équation $P_1(M) + \alpha P_2(M) = 0$. La théorie des faisceaux de plans montre qu'on est alors ramené au cas précédent.

- **6 Symétrique d'un point par rapport à une droite ou un plan**

  – On écrit que, si M' est symétrique (orthogonal) de M par rapport à $\mathcal{D}$ (ou $\mathcal{P}$), on a $\frac{1}{2}(M + M') \in \mathcal{D}$ (ou $\mathcal{P}$) et $\overrightarrow{MM'} \perp \mathcal{D}$ (ou $\mathcal{P}$).

  – Si la droite $\mathcal{D}$ est définie par un point A et un vecteur directeur $\vec{u}$, on peut aussi remarquer que

$$\overrightarrow{AM} + \overrightarrow{AM'} = 2\frac{\overrightarrow{AM}.\vec{u}}{\|\vec{u}\|^2}\vec{u}$$

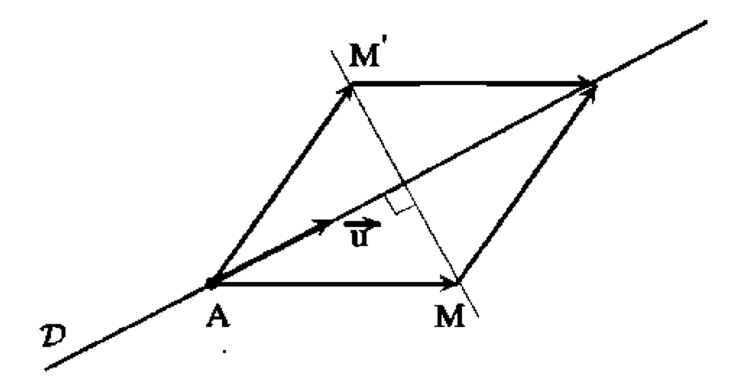


Figure **22.3** – *Symétrie par rapport à une droite ou un plan*

• **7 Perpendiculaire commune. Distance de deux droites**

$\mathcal{D}_1$ et $\mathcal{D}_2$ sont deux droites de l'espace non parallèles.

– Si les droites $\mathcal{D}_1 = (A_1, \vec{u_1})$ et $\mathcal{D}_2 = (A_2, \vec{u_2})$ sont définies par un point et un vecteur, la direction de la perpendiculaire commune est évidemment définie par le vecteur $\vec{u_1} \wedge \vec{u_2}$. Les pieds de cette perpendiculaire sont $H_1 = A_1 + \lambda_1\vec{u_1}$ et $H_2 = A_2 + \mu_2\vec{u_2}$ où (λ_1, μ_2) est point critique de $(\lambda, \mu) \mapsto d^2(A_1 + \lambda\vec{u_1}, A_2 + \mu\vec{u_2})$. La distance des deux droites est

$$\delta = \|\overrightarrow{H_1H_2}\| = \frac{|\overrightarrow{A_1A_2}.(\vec{u_1} \wedge \vec{u_2})|}{\|\vec{u_1} \wedge \vec{u_2}\|}$$

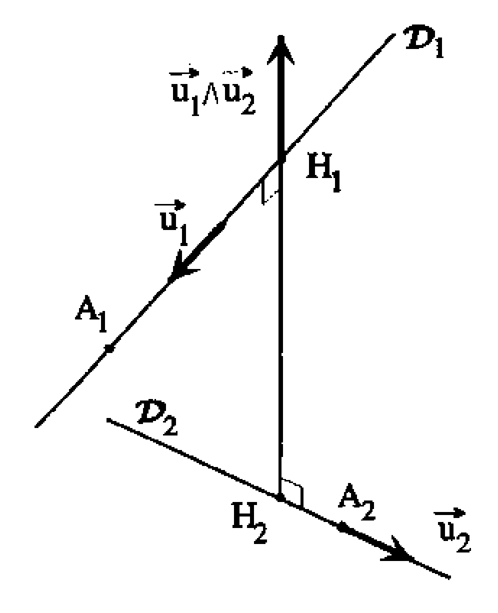


Figure **22.4** – *Perpendiculaire commune à deux droites*

– Si $\mathcal{D}_1 = \mathcal{P}_1 \cap \mathcal{Q}_1$ est intersection de deux plans d'équations d'équations $P_1(M) = 0$ et $Q_1(M) = 0$ et $\mathcal{D}_2 = \mathcal{P}_2 \cap \mathcal{Q}_2$ plans d'équations respectives $P_2(M) = 0$ et $Q_2(M) = 0$, ces équations donnent immédiatement des vecteurs $\overrightarrow{p_1}$, $\overrightarrow{q_1}$, $\overrightarrow{p_2}$, $\overrightarrow{q_2}$ normaux respectivement aux plans $\mathcal{P}_1$, $\mathcal{Q}_1$, $\mathcal{P}_2$ et $\mathcal{Q}_2$. Les vecteurs $\overrightarrow{v_1} = \overrightarrow{p_1} \wedge \overrightarrow{q_1}$ et $\overrightarrow{v_2} = \overrightarrow{p_2} \wedge \overrightarrow{q_2}$ dirigent respectivement $\mathcal{D}_1$ et $\mathcal{D}_2$, donc $\overrightarrow{w} = \overrightarrow{v_1} \wedge \overrightarrow{v_2}$ dirige la perpendiculaire commune. On peut déterminer cette dernière comme intersection de deux plans, l'un du faisceau d'arête $\mathcal{D}_1$ parallèle à $\overrightarrow{w}$ (donc son équation sera de la forme $P_1(M) + \alpha Q_1(M) = 0$ avec $(\overrightarrow{p_1} + \alpha\overrightarrow{q_1}).\overrightarrow{w} = 0$) l'autre du faisceau d'arête $\mathcal{D}_2$ vérifiant une condition similaire.
La distance des deux droites est aussi la distance de deux plans parallèles, l'un du faisceau d'arête $\mathcal{D}_1$ parallèle à $\overrightarrow{v_2}$, l'autre du faisceau d'arête $\mathcal{D}_2$ parallèle à $\overrightarrow{v_1}$. Si $ux+vy+wz+h = 0$ et $ux+vy+wz+h' = 0$ sont les équations de ces deux plans, cette distance est évidemment distance d'un point quelconque d'un des plans à l'autre plan, soit

$$\frac{|h - h'|}{\sqrt{u^2 + v^2 + w^2}}$$

- 8 **Cercle passant par trois points, sphère passant par quatre points**

On sait que, dans le plan rapporté à un repère orthonormé, l'équation du cercle de centre $\Omega(\alpha, \beta)$ et de rayon R traduit la condition

$$\|\overrightarrow{\Omega M}\|^2 = R^2$$

soit $(x-\alpha)^2 + (y-\beta)^2 = R^2$. Il s'agit donc d'une équation du second degré, dont les termes de degré 2 correspondent à une forme quadratique égale (ou proportionnelle) au carré de la norme euclidienne. Lorsque ce cercle passe par l'origine, cette équation n'a pas de terme constant et s'écrit immédiatement

$$x^2 + y^2 - 2\alpha x - 2\beta y = 0$$

Dans le cas général, l'équation du cercle passant par trois points non alignés A, B et C peut s'écrire

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_A^2+y_A^2 & x_A & y_A & 1 \\ x_B^2+y_B^2 & x_B & y_B & 1 \\ x_C^2+y_C^2 & x_C & y_C & 1 \end{vmatrix} = 0$$

(En développant par rapport à la première ligne on trouve en effet une équation de cercle, le cofacteur de x^2+y^2 n'étant pas nul). On a évidemment un résultat analogue pour la sphère passant par 4 points non coplanaires de l'espace.

- 9 **Plan osculateur à un arc**

– Si $t \mapsto M(t)(x(t), y(t), z(t))$ est un arc paramétré, le **plan osculateur** en un point $M(t_0)$ est le plan affine passant par $M(t_0)$ et dont

la direction est déterminée par les deux premiers vecteurs dérivés indépendants en t_0. En particulier, si $M(t_0)$ est **birégulier**, c'est à dire que $\{\overrightarrow{M}'(t_0), \overrightarrow{M}''(t_0)\}$ est libre, l'équation de ce plan osculateur est :

$$\begin{vmatrix} X - x(t_0) & x'(t_0) & x''(t_0) \\ Y - y(t_0) & y'(t_0) & y''(t_0) \\ X - z(t_0) & z'(t_0) & z''(t_0) \end{vmatrix} = 0$$

On trouve parfois directement l'équation de ce plan sous la forme

$$aX + bY + cZ + d = 0$$

en remarquant que $t \mapsto ax(t) + by(t) + cz(t) + d$ doit être un infiniment petit d'ordre au moins 3 au voisinage de t_0.

– On a la même caractérisation en dimension 2 pour l'équation

$$ax + by + c = 0$$

de la tangente à un arc régulier en un point de paramètre t_0. La quantité $ax(t) + by(t) + c$ doit être infiniment petit d'ordre au moins 2 en t_0. Si c'est un infiniment petit d'ordre supérieur ou égal à 3, il s'agit d'une tangente en un point d'inflexion analytique (les deux premiers vecteurs dérivés appartiennent à la direction de la droite). Si c'est de plus un infiniment petit d'ordre impair, l'arc traverse sa tangente au point de paramètre t_0 et on a effectivement un point d'inflexion géométrique.

• 10 Propriétés métriques des arcs plans

– γ est un arc plan $t \mapsto M(t)$ **régulier** de classe C^2. Un **paramétrage normal** de γ orienté dans le sens des t croissants est une fonction $t \mapsto s(t)$ telle que

$$ds = \|\overrightarrow{M}'(t)\| dt$$

s est alors un paramètre admissible pour γ puisque $t \mapsto s(t)$ est un C^2-difféomorphisme.

– **Repère de Frenet** : On note alors $\overrightarrow{T} = \frac{d\overrightarrow{M}}{ds} = \frac{\overrightarrow{M}'(t)}{\|\overrightarrow{M}'(t)\|}$. C'est le vecteur unitaire de la tangente orientée dans le sens des t croissants. Si, dans la base orthonormée directe $\{\overrightarrow{\imath}, \overrightarrow{\jmath}\}$ on a

$$\overrightarrow{T} = \cos(\alpha)\overrightarrow{\imath} + \sin(\alpha)\overrightarrow{\jmath}$$

le vecteur normal

$$\overrightarrow{N} = -\sin(\alpha)\overrightarrow{\imath} + \cos(\alpha)\overrightarrow{\jmath}$$

vérifie

$$(\widehat{\overrightarrow{T}, \overrightarrow{N}}) = \frac{\pi}{2} \bmod(2\pi)$$

et $(M(t), \overrightarrow{T}, \overrightarrow{N})$ est le **repère de Frenet** en $M(t)$.

– **Courbure** : En posant $c(s) = \dfrac{d\alpha}{ds}$, on a alors

$$\frac{d\overrightarrow{M}}{ds} = \overrightarrow{T} \ , \ \frac{d\overrightarrow{T}}{ds} = c(s)\overrightarrow{N} = \frac{d^2\overrightarrow{M}}{ds^2} \ \text{ et } \ \frac{d\overrightarrow{N}}{ds} = -c(s)\overrightarrow{T}$$

γ est **birégulier** si pour tout $t \in I$ $(\overrightarrow{M'}(t), \overrightarrow{M''}(t))$ est libre, ce qui équivaut à

$$\forall s \in J \quad c(s) \neq 0$$

On définit alors le **rayon de courbure** $R = \dfrac{1}{c(s)} = \dfrac{ds}{d\alpha}$ et le centre de courbure $C = M + R\overrightarrow{N}$. Le lieu du centre de courbure est la **développée** de γ. Le **cercle de courbure** en M, de centre C et de rayon $|R|$ est osculateur à γ en M (contact d'ordre supérieur ou égal à 3) si l'arc est de classe au moins C^3.

– **Calcul de la courbure** : On a $\overrightarrow{M'}(t) = \dfrac{ds}{dt}\overrightarrow{T}$ et donc

$$\frac{d^2\overrightarrow{M}}{dt^2} = \frac{d^2s}{dt^2}\overrightarrow{T} + \left(\frac{ds}{dt}\right)^2 c(s)\overrightarrow{N}$$

On en déduit, en notant $[\overrightarrow{x}, \overrightarrow{y}]$ le produit mixte de deux vecteurs

$$c(s) = \frac{[\overrightarrow{M'}(t), \overrightarrow{M''}(t)]}{\left(\dfrac{ds}{dt}\right)^3} = \frac{[\overrightarrow{M'}(t), \overrightarrow{M''}(t)]}{\|M'(t)\|^3}$$

– Par exemple, pour un arc défini par un paramétrage polaire

$$\overrightarrow{OM}(\theta) = r(\theta)\overrightarrow{u}(\theta)$$

on a : $\overrightarrow{M'} = r'\overrightarrow{u} + r\overrightarrow{v}$ et $\overrightarrow{M''} = (r'' - r)\overrightarrow{u} + 2r'\overrightarrow{v}$. On en déduit

$$c = \frac{r^2 + 2r'^2 - rr''}{(r^2 + r'^2)^{\frac{3}{2}}}$$

Les calculs sont parfois plus simples lorsqu'on travaille avec $q(\theta) = \dfrac{1}{r(\theta)}$ (c'est le cas pour les coniques avec foyer au pôle). Le calcul des dérivées successives du vecteur $\overrightarrow{OM}(\theta) = \dfrac{1}{q(\theta)}\overrightarrow{u}(\theta)$ donne alors

$$c = \frac{q^3(q + q'')}{(q^2 + q'^2)^{\frac{3}{2}}}$$

– Pour un arc tangent en O à l'axe Ox et orienté dans le sens des x croissants, le rayon de courbure en O vaut

$$R = \lim_{x \to 0} \frac{x^2}{2y}$$

Cette remarque permet de calculer (en utilisant un changement de repère) le rayon de courbure en un point M_0 sans nécessairement faire la rectification de γ.

– Dans le cas des arcs en dimension 3, la courbure n'est plus algébrique. Le vecteur normal principal (situé dans le plan osculateur) en un point birégulier est unitaire et défini par la relation

$$\frac{d\overrightarrow{T}}{ds} = c(s)\overrightarrow{N} \quad \text{et} \quad c(s) > 0$$

- **11** Surface d'équation $f(x,y,z)=0$

– f est une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^3$.

$$\mathcal{S} = \{M(x,y,z) \in E_3 \mid f(x,y,z) = 0\}$$

– Un point $M_0(x_0,y_0,z_0) \in \mathcal{S}$ est dit **régulier** ssi $\overrightarrow{\text{grad}}f(M_0) \neq \overrightarrow{0}$. Sur un voisinage $\mathcal{V}_0$ de M_0, $\mathcal{S}$ admet une paramétrisation :

$$\text{si } \frac{\partial f}{\partial z}(M_0) \neq 0 \quad M(x,y,z) \in \mathcal{S} \cap \mathcal{V}_0 \Longleftrightarrow z = \varphi(x,y)$$

où φ est une fonction de classe $\mathcal{C}^1$ définie sur un voisinage $\mathcal{V}$ de (x_0,y_0). $\mathcal{S} \cap \mathcal{V}_0$ est alors le support de la nappe paramétrée de classe $\mathcal{C}^1$

$$(x,y) \in \mathcal{V} \mapsto M(x,y,\varphi(x,y))$$

– En particulier, si Π_0 est le plan tangent à cette nappe en M_0, un point $P(X,Y,Z)$ appartient à Π_0 ssi

$$\overrightarrow{M_0P}.\overrightarrow{\text{grad}}f(M_0) = 0 \Longleftrightarrow \frac{\partial f}{\partial x}(M_0)(X-x_0) + \frac{\partial f}{\partial y}(M_0)(Y-y_0) + \frac{\partial f}{\partial z}(M_0)(Z-z_0) = 0$$

- **12** Nappe paramétrée régulière

– Si $\mathcal{U}$ est un ouvert de $\mathbb{R}^2$, une nappe paramétrée de classe $\mathcal{C}^1$ est une application de classe $\mathcal{C}^1 : \mathcal{U} \to E_3$

$$(u,v) \mapsto M(u,v) \text{ de coordonnées } \begin{cases} x = P(u,v) \\ y = Q(u,v) \\ z = R(u,v) \end{cases}$$

On a alors

$$\frac{\partial\overrightarrow{M}}{\partial u} = \frac{\partial P}{\partial u}\overrightarrow{\imath} + \frac{\partial Q}{\partial u}\overrightarrow{\jmath} + \frac{\partial R}{\partial u}\overrightarrow{k} \text{ et } \frac{\partial\overrightarrow{M}}{\partial v} = \frac{\partial P}{\partial v}\overrightarrow{\imath} + \frac{\partial Q}{\partial v}\overrightarrow{\jmath} + \frac{\partial R}{\partial v}\overrightarrow{k}$$

– Le point $M_0(u_0,v_0)$ est dit régulier ssi les vecteurs $\frac{\partial\overrightarrow{M}}{\partial u}(u_0,v_0)$ et $\frac{\partial\overrightarrow{M}}{\partial v}(u_0,v_0)$ sont indépendants (sinon M_0 est dit stationnaire). Le plan tangent à la nappe au point de paramètre (u_0,v_0) est alors le plan

$$\Pi_0 : \quad M_0 + \text{vect}\left(\frac{\partial\overrightarrow{M}}{\partial u}(u_0,v_0), \frac{\partial\overrightarrow{M}}{\partial v}(u_0,v_0)\right)$$

– Le vecteur $\overrightarrow{N}_0 = \dfrac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) \wedge \dfrac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)$ est normal à la nappe en M_0 et l'équation du plan Π_0 peut s'écrire :

$$P \in \Pi_0 \iff \overrightarrow{M_0P}.\left(\frac{\partial \overrightarrow{M}}{\partial u}(u_0, v_0) \wedge \frac{\partial \overrightarrow{M}}{\partial v}(u_0, v_0)\right) = 0$$

– Au voisinage d'un point régulier $M_0(x_0, y_0, z_0)$ de paramètres égaux à $(u_0, v_0) \in \mathcal{U}$, avec par exemple

$$\begin{vmatrix} \dfrac{\partial P}{\partial u} & \dfrac{\partial P}{\partial v} \\ \dfrac{\partial Q}{\partial u} & \dfrac{\partial Q}{\partial v} \end{vmatrix} (u_0, v_0) \neq 0$$

l'application $(u, v) \mapsto (x = P(u, v), y = Q(u, v))$ est un difféomorphisme local. u et v se déterminent donc en fonction de x et y par $u = S(x, y)$ et $v = T(x, y)$. Le support de la nappe possède alors, au voisinage de (u_0, v_0) une équation cartésienne :
Le point $M(u, v)$ a pour coordonnées $(x, y, z = f(x, y))$ où la fonction f, de classe $\mathcal{C}^1$ au voisinage de (x_0, y_0) est donnée par

$$f(x, y) = R(u, v) = R(S(x, y), T(x, y))$$

Il y a donc équivalence locale des définitions par paramétrage et par équation (nappes et surfaces régulières).

- **13 Equation cartésienne d'un cylindre**

C'est une équation de la forme $f(P_1(M), P_2(M)) = 0$ où P_1 et P_2 sont deux formes linéaires indépendantes. Les génératrices du cylindres sont alors parallèles à la droite d'équations $P_1(M) = 0$ et $P_2(M) = 0$.

- **14 Equation cartésienne d'un cône**

Si f est une fonction homogène de degré α c'est à dire vérifie

$$\forall \lambda \in \mathbb{R}^* \quad f(\lambda x, \lambda y, \lambda z) = |\lambda|^\alpha f(x, y, z)$$

Si P_1, P_2, P_3 sont trois formes affines indépendantes (les trois formes linéaires associées sont indépendantes), l'équation $f(P_1(M), P_2(M), P_3(M)) = 0$ est une équation de cône dont le sommet est intersection des plans d'équations $P_1(M) = 0$, $P_2(M) = 0$ et $P_3(M) = 0$.

- **15 Equation d'une surface de révolution**

Si Ω est un point de l'espace euclidien de dimension 3 et P_1 est une forme linéaire non nulle, l'équation $f(\|\overrightarrow{\Omega M}\|^2, P_1(M)) = 0$ est équation d'une surface de révolution d'axe passant par Ω et perpendiculaire au plan d'équation $P_1(M) = 0$.

Chapitre 23

Intégrales curvilignes, intégrales multiples

23-1 Formes différentielles de degré 1

23-1.1 Forme différentielle, champ de vecteurs

23-1.1.1 Forme différentielle de degré 1

Dans cette section, $\mathcal{A}_n$ est un espace affine basé sur un espace vectoriel normé $(E_n, \| \|)$ et $\mathcal{U}$ est un ouvert de $\mathcal{A}_n$. Si $\mathcal{B} = (\overrightarrow{e}_i)_{1 \leqslant i \leqslant n}$ est une base de E_n et O est un point de $\mathcal{A}_n$, on associe à $\mathcal{U}$ un ouvert U de $\mathbb{R}^n$:

$$U = \{(x_1, \dots, x_n) \in \mathbf{R}^n \mid M(x_1, \dots, x_n) \in \mathcal{U}\}$$

les coordonnées étant écrites dans le repère $\mathcal{R} = (O, \mathcal{B})$. Rappelons que l'espace dual E_n^* est l'espace des formes linéaires sur E_n. La base duale de $\mathcal{B}$ est la base de E_n^* composée des formes coordonnées dans $\mathcal{B}$.

Notations : comme c'est l'usage (*cf.* section **18-1.2.2**), si on note $(x_1, \dots, x_n)$ les coordonnées du vecteur "générique" de E_n

$$\overrightarrow{x} = \sum_{i=1}^{n} x_i \overrightarrow{e}_i$$

les formes coordonnées dans la base $\mathcal{B}$ seront notées $(dx_i)_{1 \leqslant i \leqslant n}$. Si φ est une forme linéaire sur E_n représentée par la ligne $L = (\alpha_1, \dots, \alpha_n)$ lorsqu'on travaille

dans la base $\mathcal{B}$, on aura

$$\varphi = \sum_{i=1}^{n} \alpha_i \, dx_i$$

DÉFINITION 23-1.1 *Une forme différentielle de degré* 1 *sur l'ouvert* $\mathcal{U}$ *de* $\mathcal{A}_n$ *est une application* ω *de* $\mathcal{U}$ *dans l'espace dual* E_n^*.

Si la base $\mathcal{B}$ est fixée, et $\mathcal{B}^* = (dx_i)_{1\leqslant i\leqslant n}$ est sa base duale, une forme différentielle ω de degré 1 est donc caractérisée par la donnée de n applications numériques $(P_i)_{1\leqslant i\leqslant n}$

$$P_i : \mathcal{U} \to \mathbb{R} \quad M \mapsto P_i(M)$$

telles que

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^{n} P_i(M) \, dx_i$$

On commettra souvent l'abus d'écriture consistant à confondre l'ouvert $\mathcal{U}$ et sa traduction U dans $\mathbb{R}^n$ (en identifiant le point M et le n-uplet de ses coordonnées dans $\mathcal{R} = (O, \mathcal{B})$), et on notera

$$\omega(M) = \sum_{i=1}^{n} P_i(x_1, \dots, x_n) \, dx_i$$

On a alors, pour $M(x_1, \dots, x_n) \in \mathcal{U}$ et $h = \sum_{i=1}^{n} h_i \vec{e}_i \in E_n$

$$\omega(M)(h) = \sum_{i=1}^{n} P_i(x_1, \dots, x_n) \, h_i$$

Lorsqu'on change l'origine du repère $\mathcal{R}$, l'écriture de ω ne change pas (on effectue simplement une translation sur les arguments des fonctions P_i si on les considère comme étant définies sur un ouvert de $\mathbb{R}^n$). Lorsqu'on effectue un changement de base, les fonctions P_i se transforment comme composantes d'une forme linéaire :

PROPOSITION 23-1.2 Soient $\mathcal{B} = (\vec{e}_i)_{1\leqslant i\leqslant n}$ et $\mathcal{B}' = (\vec{u}_i)_{1\leqslant i\leqslant n}$ deux bases de E_n et $P = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, les bases duales étant notées

$$\mathcal{B}^* = (dx_i)_{1\leqslant i\leqslant n} \quad \textbf{et} \quad \mathcal{B}'^* = (dy_i)_{1\leqslant i\leqslant n}$$

Si une forme différentielle est définie sur un ouvert de $\mathcal{A}_n$ par

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^{n} P_i(M) \, dx_i = \sum_{i=1}^{n} Q_i(M) \, dy_i$$

on a

$$\forall M \in \mathcal{U} \quad \begin{pmatrix} Q_1(M) \\ \vdots \\ Q_n(M) \end{pmatrix} = {}^tP \begin{pmatrix} P_1(M) \\ \vdots \\ P_n(M) \end{pmatrix}$$

Démonstration : C'est simplement la formule de changement de base duale (*cf.* exercice **4-2.6**). La formule de changement de coordonnées lorsqu'on passe de $\mathcal{B}$ à $\mathcal{B}'$ peut s'écrire symboliquement

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = P \begin{pmatrix} dy_1 \\ \vdots \\ dy_n \end{pmatrix}$$

L'expression des $(Q_j)_{1 \leqslant j \leqslant n}$ en fonction des $(P_i)_{1 \leqslant i \leqslant n}$ en résulte. ■

DÉFINITION 23-1.3 *Une forme différentielle de degré* 1 *définie sur* $\mathcal{U}$ *par*

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^{n} P_i(M)\, dx_i$$

est de classe C^k *(*$k \in \mathbb{N}$*) si les applications* $P_i : \mathcal{U} \to \mathbb{R}$ *le sont. Ceci ne dépend évidemment pas du repère choisi dans* $\mathcal{A}_n$.

Comme E_n^* est un espace vectoriel, on peut évidemment munir l'ensemble des formes différentielles de degré 1 sur $\mathcal{U}$ d'une structure d'espace vectoriel : si ω et ω' sont de telles formes et si $\alpha \in \mathbb{R}$, on aura simplement par définition

$$(\omega + \alpha\omega')(M) = \omega(M) + \alpha\,\omega'(M)$$

L'ensemble des formes de classe C^k est évidemment un sous-espace vectoriel.

EXEMPLE 23-1.4 Si $f : \mathcal{U} \to \mathbb{R}$ est une fonction numérique de classe C^1 définie sur $\mathcal{U}$, on lui associe de manière naturelle la forme différentielle $\omega = df$, définie par

$$\forall M \in \mathcal{U} \quad \omega(M) = df_M = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(M)\, dx_i$$

Si f est de classe C^{k+1}, la forme ω est évidemment de classe C^k.

23-1.1.2 Cas d'un espace euclidien : champ de vecteurs

Si l'espace affine $\mathcal{A}_n$ est euclidien, on sait que, pour toute forme linéaire φ sur l'espace E_n, il existe un unique vecteur $\overrightarrow{z} \in E_n$ tel que

$$\varphi = (\overrightarrow{z}, \bullet)$$

(identification de l'espace et de son dual à l'aide du produit scalaire). Si

$$\omega : \mathcal{U} \to E_n^*$$

est une forme différentielle de classe C^k, on peut lui associer une application $\overrightarrow{V} : \mathcal{U} \to E_n$ par

$$\forall M \in \mathcal{U} \quad \omega(M) = \left(\overrightarrow{V}(M), \bullet\right)$$

Il est clair que $\overrightarrow{V}$ sera également de classe $\mathcal{C}^k$. On dit que $\overrightarrow{V}$ est le champ de vecteur dans l'ouvert $\mathcal{U}$ associé à la forme différentielle ω. Réciproquement, tout champ de vecteurs de classe $\mathcal{C}^k$ dans $\mathcal{U}$ permet de définir une forme différentielle de degré 1 et de classe $\mathcal{C}^k$ dans $\mathcal{U}$. L'ensemble des champs de vecteurs de classe $\mathcal{C}^k$ sur $\mathcal{U}$ est un espace vectoriel pour les opérations évidentes de somme et multiplication par un scalaire.

Si on travaille dans un repère $\mathcal{R} = (O, \mathcal{B})$ orthonormé, le passage d'une notion à l'autre est simple : si $\mathcal{B} = (\overrightarrow{e}_i)_{1\leqslant i\leqslant n}$ est orthonormale et si

$$\omega(M) = \sum_{i=1}^{n} P_i(M)\ dx_i$$

on aura simplement

$$\overrightarrow{V}(M) = \sum_{i=1}^{n} P_i(M)\ \overrightarrow{e_i}$$

(dans une base non orthonormale, il faudrait faire intervenir la matrice du produit scalaire pour calculer les composantes de $\overrightarrow{V}(M)$ en fonction des $(P_i(M))_{1\leqslant i\leqslant n}$).

Lorsqu'on effectue un changement de bases orthonormales, la formule de la proposition **23-1**.2 peut d'ailleurs être interprétée comme formule de changement de coordonnées d'un vecteur : si P est matrice de changement de bases orthonormales, on a ${}^tP = P^{-1}$.

23-1.1.3 Intégrale, circulation

DÉFINITION 23-1.5 *Soit ω une forme différentielle de degré 1 continue sur un ouvert $\mathcal{U}$ de $\mathcal{A}_n$ et soit*

$$\gamma : [a, b] \ni t \mapsto M(t) \in \mathcal{U}$$

un arc paramétré de classe C^1 dont le support est inclus dans $\mathcal{U}$. L'intégrale de ω le long de γ est

$$\int_\gamma \omega \underset{\text{déf}}{=} \int_a^b \omega(M(t))\left(\overrightarrow{M}'(t)\right)\ dt$$

Si on fixe un repère $\mathcal{R} = (O, \mathcal{B})$ de $\mathcal{A}_n$ l'arc paramétré γ est déterminé par la donnée des coordonnées $(x_1(t), \ldots, x_n(t))$ de $M(t)$ dans $\mathcal{R}$ (donc par un n-uplet de fonctions numériques de classe C^1). Si ω est donnée par

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^{n} P_i(M)\ dx_i$$

on aura alors

$$\int_\gamma \omega = \int_a^b \sum_{i=1}^{n} P_i(M(t))\ x_i'(t)\ dt$$

ce qu'on pourra encore écrire, en considérant que chaque application P_i est définie sur l'ouvert U de $\mathbb{R}^n$

$$\int_\gamma \omega = \int_a^b \sum_{i=1}^n P_i\left(x_1(t), \ldots, x_n(t)\right) x_i'(t)\, dt$$

Cette intégrale existe bien, puisque la fonction considérée est continue sur $[a,b]$. Etant donné son expression, on utilisera aussi le symbole

$$\int_\gamma \omega = \int_\gamma \sum_{i=1}^n P\left(x_1, \ldots, x_n\right) dx_i$$

PROPOSITION 23-1.6 L'intégrale d'une forme différentielle ω continue de degré 1 le long d'un arc γ de classe C^1 est invariante par changement de paramétrage admissible strictement croissant. Elle est multipliée par -1 si on intègre le long d'un arc se déduisant de γ par changement de paramétrage admissible strictement décroissant.

Démonstration : Soient

$$\gamma : [a,b] \ni t \mapsto M(t) \text{ et } \gamma_1 : [\alpha,\beta] : u \mapsto P(u)$$

deux arcs C^1-équivalents tracés dans $\mathcal{U}$. Il existe donc un C^1-difféomorphisme $\theta : [a,b] \to [\alpha,\beta]$ tel que

$$M = P \circ \theta$$

On a donc

$$\int_\gamma \omega = \int_a^b \omega(M(t))\left(\overrightarrow{M}'(t)\right) dt$$
$$= \int_a^b \omega(P(\theta(t)))\left(\overrightarrow{P}'(\theta(t))\right) \theta'(t)\, dt$$

Comme θ est de classe C^1 sur $[a,b]$ et comme la fonction

$$u \mapsto \omega(P(u))\left(\overrightarrow{P}'(u)\right)$$

est continue sur $[\alpha,\beta]$, la formule de changement de variable donne

$$\int_\gamma \omega = \int_{\theta(a)}^{\theta(b)} \omega(P(u))\left(\overrightarrow{P}'(u)\right) du$$

Si θ est strictement croissant, on a $\theta(a) = \alpha$ et $\theta(b) = \beta$, et on retrouve l'intégrale de ω sur γ_1. On obtient l'opposé de cette intégrale si θ décroît. ■

REMARQUE 23-1.7 Si γ est un arc paramétré $[a,b] \to \mathcal{U}$ supposé **continu et de classe C^1 par morceaux**, on pourra définir l'intégrale de ω sur γ en utilisant, comme d'habitude, une subdivision de $[a,b]$ adaptée à γ, et en coupant l'intégrale par la relation de Chasles.

Les propriétés qui suivent découlent immédiatement de la définition :

- Si ω_1 et ω_2 sont deux formes différentielles de degré 1 continues sur $\mathcal{U}$ et si γ est un arc paramétré continu et C^1 par morceaux tracé dans $\mathcal{U}$

$$\forall \lambda \in \mathbb{R} \quad \int_\gamma (\omega_1 + \lambda\,\omega_2) = \int_\gamma \omega_1 + \lambda \int_\gamma \omega_2$$

- Si $\gamma_1 : [a,b] \ni t \mapsto M(t)$ et $\gamma_2 : [\alpha,\beta] \ni t \mapsto N(t)$ sont deux arcs continus et C^1 par morceaux tracés dans $\mathcal{U}$ avec $M(b) = N(\alpha)$, on peut "juxtaposer" γ_1 et γ_2 en un arc γ continu et C^1 par morceaux défini par

$$\gamma : [a, b+\beta-\alpha] \ni t \mapsto P(t)$$

avec $P(t) = M(t)$ si $t \in [a,b]$ et $P(t) = N(t+\alpha-b)$ sinon (on peut aussi envisager tout paramétrage équivalent respectant l'orientation). On a alors

$$\int_\gamma \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega$$

Lorsque l'espace est euclidien, on préfère souvent faire intervenir le champ de vecteur associé à une forme différentielle de degré 1. L'intégrale porte alors le nom de circulation du champ :

DÉFINITION 23-1.8 *La circulation d'un champ $\overrightarrow{V}$ de vecteurs continu dans $\mathcal{U}$ le long de l'arc paramétré $\gamma : [a,b] \to M(t) \in \mathcal{U}$ continu et C^1 par morceaux est*

$$\int_\gamma \overrightarrow{V}(M).d\overrightarrow{M} \underset{déf}{=} \int_a^b \overrightarrow{V}(M(t)).\overrightarrow{M}'(t)\,dt$$

Cette circulation possède évidemment les mêmes propriétés que l'intégrale des formes différentielles de degré 1.

23-1.2 Forme exacte, forme fermée

23-1.2.1 Définitions

DÉFINITION 23-1.9 *Soit k un entier naturel. Une forme différentielle de classe C^k $\omega : \mathcal{U} \to E_n^*$ est exacte dans $\mathcal{U}$ si et seulement s'il existe une fonction numérique $f : \mathcal{U} \to \mathbb{R}$ de classe C^{k+1} telle que*

$$\forall M \in \mathcal{U} \quad \omega(M) = df_M$$

Si ω est définie à l'aide du choix d'une base $\mathcal{B} = (\overrightarrow{e}_i)_{1 \leqslant i \leqslant n}$ par

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^n P_i(M)\,dx_i$$

cela signifie qu'il existe $f : \mathcal{U} \to \mathbb{R}$ admettant des dérivées partielles d'ordre 1 (dans la base $\mathcal{B}$) par rapport à chaque variable vérifiant

$$\forall i \quad \forall M \in \mathcal{U} \quad \frac{\partial f}{\partial x_i}(M) = P_i(M)$$

(l'hypothèse faite sur ω montre qu'alors f est nécessairement de classe C^{k+1}).

DÉFINITION 23-1.10 *Si ω est une forme différentielle de classe C^k exacte dans $\mathcal{U}$, toute fonction $f : \mathcal{U} \to \mathbb{R}$ vérifiant $df = \omega$ sur $\mathcal{U}$ est appelée primitive de ω dans $\mathcal{U}$.*

PROPOSITION 23-1.11 Si $\mathcal{U}$ est connexe, deux primitives d'une même forme différentielle ω exacte dans $\mathcal{U}$ diffèrent d'une constante.

Démonstration : Si f et g sont deux primitives de ω, on a

$$\forall M \in \mathcal{U} \quad d(f-g)_M = 0_{E_n^*}$$

et on sait que ceci entraîne que $f-g$ est constante sur $\mathcal{U}$ si cet ouvert est connexe (sinon, $f-g$ est constante sur chaque composante connexe de $\mathcal{U}$). ■

Si **l'espace est euclidien**, une forme différentielle ω est exacte si et seulement si le champ de vecteurs $\vec{V}$ correspondant est un **champ de gradient**

$$\forall M \in \mathcal{U} \quad \vec{V}(M) = \overrightarrow{\text{grad}}\, f(M)$$

(avec toujours f de classe C^{k+1} si le champ est de classe C^k). Toute fonction f vérifiant cette égalité est un potentiel (scalaire) du champ $\vec{V}$. On dit aussi que $\vec{V}$ dérive du potentiel f.

23-1.2.2 Condition nécessaire : forme fermée

On suppose que ω est une forme différentielle de classe C^k (avec $k \geqslant 1$) qui s'exprime dans une base $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$ par

$$\omega(M) = \sum_{i=1}^{n} P_i(M)\, dx_i$$

Si ω est exacte et si f en est une primitive, on a

$$\forall i \quad \forall M \in \mathcal{U} \quad \frac{\partial f}{\partial x_i}(M) = P_i(M)$$

Comme f est au moins de classe C^2 (puisque les fonctions P_i sont de classe C^k, avec $k \geqslant 1$), on obtient pour $i \neq j$ dans $\{1, \dots, n\}$

$$\forall M \in \mathcal{U} \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(M) = \frac{\partial P_i}{\partial x_j}(M) \quad \text{et} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(M) = \frac{\partial P_j}{\partial x_i}(M)$$

Comme f vérifie le théorème de Schwarz, on est amené à la définition suivante

DÉFINITION 23-1.12 *Une forme différentielle ω de classe C^k (avec $k \geqslant 1$) définie sur $\mathcal{U}$ par*

$$\forall M \in \mathcal{U} \quad \omega(M) = \sum_{i=1}^{n} P_i(M)\, dx_i$$

est fermée dans $\mathcal{U}$ si et seulement si

$$\forall M \in \mathcal{U}\ \forall i \neq j \in \{1, \dots, n\} \quad \frac{\partial P_i}{\partial x_j}(M) = \frac{\partial P_j}{\partial x_i}(M) \quad (*)$$

PROPOSITION 23-1.13 Toute forme différentielle exacte ω de classe C^k (avec $k \geqslant 1$) définie sur $\mathcal{U}$ est fermée dans $\mathcal{U}$.

Les conditions $(*)$ sont donc **nécessaires** pour que ω soit **exacte**. Nous verrons (théorème **23-1**.19) que ces conditions sont suffisantes avec une hypothèse supplémentaire sur la géométrie de l'ouvert $\mathcal{U}$, vérifiée notamment par les boules ouvertes. Il en résultera qu'une forme **fermée** dans un ouvert est toujours **localement exacte**.

REMARQUE 23-1.14 La définition **23-1**.12 semble dépendre du choix de la base $\mathcal{B}$. On pourrait vérifier qu'il n'en est rien en utilisant les formules de changement de base (proposition **23-1**.2) dans laquelle il faut utiliser deux systèmes de coordonnées du même point $M \in \mathcal{U}$. Ce caractère intrinsèque sera en fait établi par le théorème **23-1**.19.

23-1.2.3 Intégrale d'une forme exacte

PROPOSITION 23-1.15 Soit ω une forme différentielle continue et exacte dans un ouvert $\mathcal{U}$, et f une primitive de ω dans $\mathcal{U}$. Si $\gamma : [a,b] \ni t \mapsto M(t)$ est un arc continu et C^1 par morceaux tracé dans $\mathcal{U}$, on a

$$\int_\gamma \omega = f(M(b)) - f(M(a))$$

Cette intégrale ne dépend donc que de l'origine et de l'extrémité de l'arc γ.

Démonstration : On a en effet, par définition de l'intégrale

$$\int_\gamma \omega = \int_a^b df_{M(t)}\left(\overrightarrow{M}'(t)\right) dt = \int_a^b \frac{d}{dt}\left(f(M(t))\right) dt$$

d'après la formule de dérivation d'une fonction composée. Comme f est de classe C^1 dans $\mathcal{U}$, l'application $f \circ M$ est continue et C^1 par morceaux sur $[a,b]$, ce qui donne le résultat par application de la formule fondamentale du calcul intégral. ■

COROLLAIRE 23-1.16 L'intégrale d'une forme différentielle exacte le long d'un arc fermé continu et C^1 par morceaux tracé dans $\mathcal{U}$ est nulle.

COROLLAIRE 23-1.17 On suppose l'espace euclidien. Si $f : \mathcal{U} \to \mathbb{R}$ est une fonction de classe C^1 et si $\gamma : [a,b] \ni t \mapsto M(t)$ est un arc continu et C^1 par morceaux tracé dans $\mathcal{U}$, on a

$$\int_\gamma \overrightarrow{\mathrm{grad}}\, f(M) . d\overrightarrow{M} = f(M(b)) - f(M(a))$$

En particulier, la circulation d'un champ de gradient le long d'un arc fermé est nulle : on dit qu'un champ de gradient est à circulation conservative.

Ces résultats nous permettent de montrer que la proposition **23-1**.13 ne donne pas une condition suffisante pour qu'une forme différentielle de classe C^1 soit exacte : dans $\mathbb{R}^2$, considérons l'ouvert $\mathcal{U} = \mathbb{R}^2 - \{(0,0)\}$ et la forme différentielle de classe C^∞ définie sur $\mathcal{U}$ par

$$\omega = -\frac{y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = P(x,y)\,dx + Q(x,y)\,dy$$

Cette forme est fermée dans $\mathcal{U}$, car l'égalité

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

est vérifiée en tout point de $\mathcal{U}$. L'arc $\gamma : [0, 2\pi] \ni t \mapsto (\cos t, \sin t)$ est fermé, C^∞ et son support est tracé dans $\mathcal{U}$. On a

$$\int_\gamma \omega = \int_0^{2\pi} \left[-\frac{\sin t}{\cos^2 t + \sin^2 t}(-\sin t) + \frac{\cos t}{\cos^2 t + \sin^2 t}(\cos t)\right] dt = 2\pi$$

Comme cette intégrale n'est pas nulle, ω ne peut être exacte.

23-1.2.4 Théorème de Poincaré

DÉFINITION 23-1.18 *Soit un ouvert $\mathcal{U}$ d'un espace affine et M_0 un point de $\mathcal{U}$. On dit que $\mathcal{U}$ est étoilé par rapport à M_0 si et seulement si*

$$\forall M \in \mathcal{U} \quad \text{le segment } [M_0, M] \text{ est inclus dans } \mathcal{U}$$

Un ouvert $\mathcal{U}$ sera dit étoilé si et seulement si

$$\exists M_0 \in \mathcal{U} \qquad \mathcal{U} \text{ est étoilé par rapport à } M_0$$

Par exemple, un ouvert convexe est un ouvert étoilé par rapport à chacun de ses points. Dans $\mathbb{R}^2$, l'ouvert $\mathbb{R}^2 - (\mathbb{R}^- \times \{0\})$ est étoilé alors que $\mathbb{R}^2 - \{(0,0)\}$ ne l'est pas. Un ouvert étoilé est évidemment connexe par arcs.

THÉORÈME 23-1.19 Si $\mathcal{U}$ est un ouvert étoilé, toute forme différentielle de degré 1 de classe C^1 et fermée dans $\mathcal{U}$ est exacte. En particulier, dans toute boule ouverte

$$\omega \textbf{ fermée } \Rightarrow \omega \textbf{ exacte}$$

Démonstration : On choisit un point $O \in \mathcal{U}$ tel que $\mathcal{U}$ soit étoilé par rapport à O, et on prend un repère $\mathcal{R} = (O, \mathcal{B})$ où $\mathcal{B} = (\vec{e}_i)_{1 \leqslant i \leqslant n}$, dans lequel ω s'écrit

$$\omega(M) = \sum_{i=1}^{n} P_i(M)\,dx_i$$

les P_i étant de classe C^1 et vérifiant

$$\forall i \neq j \in \{1, \ldots, n\} \quad \frac{\partial P_i}{\partial x_j}(M) = \frac{\partial P_j}{\partial x_i}(M)$$

Si M est un point quelconque de $\mathcal{U}$, l'arc de classe $\mathcal{C}^\infty$

$$\gamma_M : [0,1] \ni t \mapsto M(t) = O + t\overrightarrow{OM}$$

a son support tracé dans $\mathcal{U}$.

Analyse : Si une primitive f de ω existe, on aura nécessairement, d'après la proposition **23-1**.15

$$\forall M \in \mathcal{U} \quad f(M) - f(O) = \int_{\gamma_M} \omega$$

Synthèse : Comme la primitive est définie à une constante près, nous définirons f par cette formule, en imposant $f(O) = 0$ (par exemple). En identifiant M avec le n-uplet $(x_1, \dots, x_n)$ de ses coordonnées dans $\mathcal{R}$, nous aurons donc

$$f(x_1, \dots, x_n) = \int_0^1 \sum_{i=1}^n P_i(M(t))\, x_i\, dt$$

puisque $\overrightarrow{M}'(t) = \overrightarrow{OM}$. Comme les coordonnées de $M(t)$ sont évidemment égales à $(tx_1, \dots, tx_n)$, on a

$$f(x_1, \dots, x_n) = \int_0^1 \sum_{i=1}^n P_i(tx_1, \dots, tx_n)\, x_i\, dt$$

Si on fixe $(x_2, \dots, x_n)$, on considère la fonction (définie au voisinage de x_1, donc au moins sur un intervalle ouvert de la forme $]x_1 - \alpha, x_1 + \alpha[$)

$$x \mapsto f(x, x_2, \dots, x_n) = \\ \int_0^1 \left[P_1(tx, tx_2, \dots, tx_n)\, x + \sum_{i=2}^n P_i(tx, tx_2, \dots, tx_n)\, x_i \right] dt$$

On peut lui appliquer le théorème de dérivation sous le signe d'intégration (**10-2**.2) puisque l'application

$$\varphi : [0,1] \times]x_1 - \alpha, x_1 + \alpha[\to \mathbb{R} \\ \varphi(t,x) = P_1(tx, tx_2, \dots, tx_n)\, x + \sum_{i=2}^n P_i(tx, tx_2, \dots, tx_n)\, x_i$$

est continue sur $[0,1] \times]x_1 - \alpha, x_1 + \alpha[$ et possède une dérivée partielle par rapport à la seconde variable

$$\frac{\partial \varphi}{\partial x}(t,x) = P_1(tx, tx_2, \dots, tx_n) \\ + tx\, \frac{\partial P_1}{\partial x_1}(tx, tx_2, \dots, tx_n) + \sum_{i=2}^n tx_i\, \frac{\partial P_i}{\partial x_1}(tx, tx_2, \dots, tx_n)$$

la fonction $\dfrac{\partial\varphi}{\partial x}$ étant aussi continue sur $[0,1] \times]x_1 - \alpha, x_1 + \alpha[$ (le calcul de dérivée est justifié par le fait que les fonctions P_i sont de classe $\mathcal{C}^1$). Comme ω est fermée, cette dérivée s'écrit aussi

$$\begin{aligned}\frac{\partial\varphi}{\partial x}(t,x) &= P_1(tx, tx_2, \dots, tx_n)\\ &\quad + tx\,\frac{\partial P_1}{\partial x_1}(tx, tx_2, \dots, tx_n) + \sum_{i=2}^{n} tx_i\,\frac{\partial P_1}{\partial x_i}(tx, tx_2, \dots, tx_n)\end{aligned}$$

soit finalement

$$\frac{\partial\varphi}{\partial x}(t,x) = \frac{d}{dt}\left[tx\,P_1(tx, tx_2, \dots, tx_n)\right]$$

On en déduit que f admet en M une dérivée partielle par rapport à la première variable

$$\begin{aligned}\frac{\partial f}{\partial x_1}(M) &= \int_0^1 \frac{d}{dt}\left[tx\,P_1(tx, tx_2, \dots, tx_n)\right] dt\\ &= \left[tx\,P_1(tx, tx_2, \dots, tx_n)\right]_0^1 = P_1(M)\end{aligned}$$

On calculerait de même les dérivées partielles par rapport aux autres variables. Si les P_i sont supposées de classe $\mathcal{C}^k$, il en résulte que f est de classe $\mathcal{C}^{k+1}$ sur $\mathcal{U}$, avec

$$\forall M \in \mathcal{U} \quad df_M = \omega(M)$$

et ω est bien exacte. ■

COROLLAIRE 23-1.20 L'espace étant euclidien, si $\vec{V}$ est un champ vectoriel $\mathcal{C}^1$ dans $\mathcal{U}$ étoilé s'exprimant dans une base orthonormale $(\vec{e}_i)_{1\leqslant i\leqslant n}$ par

$$\forall M \in \mathcal{U} \quad \vec{V}(M) = \sum_{i=1}^{n} P_i(M)\,\vec{e}_i$$

$\vec{V}$ dérive d'un potentiel si et seulement si

$$\forall i \neq j \in \{1, \dots, n\} \quad \frac{\partial P_i}{\partial x_j}(M) = \frac{\partial P_j}{\partial x_i}(M)$$

En particulier, dans toute boule ouverte, ces conditions caractérisent les champs de gradient (de classe $\mathcal{C}^1$).

EXEMPLE 23-1.21 Nous avons vu à la section **23-1.2.3** que la forme différentielle

$$\omega = -\frac{y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy$$

était fermée mais non exacte dans $\mathbb{R}^2 - \{(0,0)\}$. En restriction à tout ouvert étoilé inclus dans $\mathbb{R}^2 - \{(0,0)\}$, elle est exacte. Par exemple, sur $\mathbb{R}^2 - (\mathbb{R}^- \times \{0\})$, la fonction θ définie par

$$\theta(x,y) = 2\arctan\frac{y}{x+\sqrt{x^2+y^2}}$$

(détermination C^∞ dans cet ouvert de l'argument du complexe $x+iy$) est une primitive de ω. On comprend ainsi pourquoi l'intégrale de ω le long du cercle unité orienté positivement vaut 2π.

Dans la pratique, lorsqu'une forme différentielle est fermée, on en détermine une primitive par "intégrations successives" :

EXERCICE 23-1.22 On considère le sous-ensemble de $\mathbb{R}^3$

$$\mathcal{D} = \left\{(x,y,z)\in\mathbb{R}^3 \mid x^2-y^2\neq 0\right\}$$

Montrer que $\mathcal{D}$ est réunion de quatre ouverts convexes, et que la forme différentielle définie sur $\mathcal{D}$ par

$$\omega = \left[\frac{x^6-2x^4y^2+x^2y^4-4xzy^2}{(x^2-y^2)^2}\right]dx + \frac{4x^2yz}{(x^2-y^2)^2}\,dy + \frac{2x^2}{x^2-y^2}\,dz$$

est exacte. En donner une primitive.

Comme $\mathcal{D}$ est réunion de quatre ouverts où le théorème de Poincaré s'applique, on peut prouver l'existence de primitives (sans en donner une explicitement) en montrant que ω est fermée. Si

$$\omega = P\,dx + Q\,dy + R\,dz$$

on a bien

$$\frac{\partial P}{\partial y} = -8xyz\frac{x^2+y^2}{(x^2-y^2)^3} = \frac{\partial Q}{\partial x},\ \frac{\partial Q}{\partial z} = \frac{4x^2y}{(x^2-y^2)^2} = \frac{\partial R}{\partial y} \text{ et } \frac{\partial R}{\partial x} = -\frac{4xy^2}{(x^2-y^2)^2} = \frac{\partial P}{\partial z}$$

Pour expliciter une primitive, on cherche d'abord une fonction f de classe C^∞ dans $\mathcal{D}$ avec

$$\frac{\partial f}{\partial z} = \frac{2x^2}{x^2-y^2}$$

En "intégrant par rapport à z", on voit que ceci peut s'écrire

$$\frac{\partial}{\partial z}\left[f(x,y,z) - \frac{2x^2z}{x^2-y^2}\right] = 0$$

Ceci amène évidemment à chercher la fonction f sous la forme[1]

$$f(x,y,z) = \frac{2x^2z}{x^2-y^2} + \varphi(x,y)$$

où φ est une fonction C^∞ dans l'ouvert $\left\{(x,y)\in\mathbb{R}^2 \mid x^2\neq y^2\right\}$. Les conditions

$$\frac{\partial f}{\partial x} = P \text{ et } \frac{\partial f}{\partial y} = Q$$

[1]On est assuré localement d'une telle écriture pour f. Voir la remarque **18-3**.6.

donnent alors

$$\frac{\partial \varphi}{\partial x} = x^2 \quad \text{et} \quad \frac{\partial \varphi}{\partial y} = 0$$

On voit ainsi que l'on peut choisir

$$f(x,y,z) = \frac{x^3}{3} + \frac{2x^2 z}{x^2 - y^2} = \frac{x^5 - x^3 y^2 + 6x^2 z}{3(x^2 - y^2)}$$

On notera que l'ensemble des primitives de ω dans $\mathcal{D}$ est ici un espace affine de dimension 4.

EXERCICE 23-1.23 On rappelle que, dans un ouvert connexe d'un espace affine euclidien, on peut toujours relier deux points quelconques par un chemin polygonal (voir exercice **7-7**.22). En s'inspirant de la démonstration du théorème **23-1**.19, montrer qu'un champ de vecteurs continu dans un ouvert connexe est un champ de gradient si et seulement s'il est à circulation conservative.

23-1.3 Rotationnel, divergence, laplacien

On se place dans un espace affine euclidien orienté de dimension 3, ramené à un repère orthonormé direct $\left(O, \vec{I}, \vec{J}, \vec{K}\right)$. On se donne un champ de vecteur $\vec{V}$ de classe C^1 dans un ouvert $\mathcal{U}$

$$\vec{V}(M) = P(M)\vec{I} + Q(M)\vec{J} + R(M)\vec{K}$$

Le théorème de Poincaré nous montre que $\vec{V}$ est localement un champ de gradient (globalement dans $\mathcal{U}$ si cet ouvert est étoilé) si et seulement si le champ vectoriel (continu) $\vec{W}$ défini dans $\mathcal{U}$ par

$$\vec{W}(M) = \left(\frac{\partial R}{\partial y}(M) - \frac{\partial Q}{\partial z}(M)\right)\vec{I} + \left(\frac{\partial P}{\partial z}(M) - \frac{\partial R}{\partial y}(M)\right)\vec{J} + \left(\frac{\partial Q}{\partial x}(M) - \frac{\partial P}{\partial y}(M)\right)\vec{K}$$

est identiquement nul. Il est aussi défini par

$$\vec{W}(M) = \vec{I} \wedge \frac{\partial \vec{V}}{\partial x} + \vec{J} \wedge \frac{\partial \vec{V}}{\partial y} + \vec{K} \wedge \frac{\partial \vec{V}}{\partial z}$$

Ce champ, défini lorsque $\vec{V}$ est de classe C^1, ne dépend pas du repère orthonormal direct utilisé pour sa définition : en effet, si on considère l'application

$$\vec{V} : \mathcal{U} \to E_3$$

et si on choisit un point $M_0 \in \mathcal{U}$ arbitraire, l'endomorphisme $d\vec{V}_{M_0} \in \mathcal{L}(E_3)$ est

représenté, dans la base $\mathcal{B} = \left(\vec{I}, \vec{J}, \vec{K}\right)$, par la matrice jacobienne

$$J^{\mathcal{B}}\left(\vec{V}\right)(M_0) = \begin{pmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y} & \dfrac{\partial P}{\partial z} \\ \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y} & \dfrac{\partial Q}{\partial z} \\ \dfrac{\partial R}{\partial x} & \dfrac{\partial R}{\partial y} & \dfrac{\partial R}{\partial z} \end{pmatrix}(M_0)$$

Si la base est orthonormale, l'adjoint de $d\vec{V}_{M_0}$ (endomorphisme défini de manière intrinsèque) est représenté dans $\mathcal{B}$ par la transposée de cette matrice, et la différence $d\vec{V}_{M_0} - \left(d\vec{V}_{M_0}\right)^*$ (qui représente deux fois la partie antisymétrique de $d\vec{V}_{M_0}$) a pour matrice

$$J^{\mathcal{B}}\left(\vec{V}\right)(M_0) - {}^tJ^{\mathcal{B}}\left(\vec{V}\right)(M_0) = \begin{pmatrix} 0 & \dfrac{\partial P}{\partial y} - \dfrac{\partial Q}{\partial x} & \dfrac{\partial P}{\partial z} - \dfrac{\partial R}{\partial x} \\ \dfrac{\partial Q}{\partial x} - \dfrac{\partial P}{\partial y} & 0 & \dfrac{\partial Q}{\partial z} - \dfrac{\partial R}{\partial y} \\ \dfrac{\partial R}{\partial x} - \dfrac{\partial P}{\partial z} & \dfrac{\partial R}{\partial y} - \dfrac{\partial Q}{\partial z} & 0 \end{pmatrix}(M_0)$$

Si $\left(\vec{I}, \vec{J}, \vec{K}\right)$ est orthonormale directe, cette matrice antisymétrique est interprétée comme matrice du produit vectoriel par $\vec{W}(M_0)$:

$$d\vec{V}_{M_0} - \left(d\vec{V}_{M_0}\right)^* = \vec{W}(M_0) \wedge \bullet$$

Le champ $\vec{W}$ est appelé rotationnel de $\vec{V}$, il ne dépend donc effectivement pas de la base orthonormale directe utilisée.

PROPOSITION 23-1.24 Si $\vec{V}$ est un champ de vecteurs de classe C^1 exprimé dans un repère orthonormal direct par

$$\vec{V}(M) = P(M)\vec{I} + Q(M)\vec{J} + R(M)\vec{K}$$

son rotationnel défini par

$$\operatorname{rot}\vec{V}(M) = \left(\frac{\partial R}{\partial y}(M) - \frac{\partial Q}{\partial z}(M)\right)\vec{I} + \left(\frac{\partial P}{\partial z}(M) - \frac{\partial R}{\partial y}(M)\right)\vec{J} + \left(\frac{\partial Q}{\partial x}(M) - \frac{\partial P}{\partial y}(M)\right)\vec{K}$$

ne dépend pas de la base orthonormale (directe) choisie. Dans un ouvert étoilé, un champ de classe C^1 est un champ de gradient si et seulement si son rotationnel est identiquement nul.

Le **rotationnel,** comme le **gradient,** a donc une définition intrinsèque, ce qui en fait un outil pour exprimer certaines lois de la Physique. Il en est de même de la **divergence d'un champ vectoriel** de classe C^1. Si

$$\vec{V}(M) = P(M)\,\vec{I} + Q(M)\,\vec{J} + R(M)\,\vec{K}$$

est exprimé dans une base quelconque,

$$\operatorname{div}\vec{V}(M) = \frac{\partial P}{\partial x}(M) + \frac{\partial Q}{\partial y}(M) + \frac{\partial R}{\partial z}(M)$$

n'est autre que la trace de la matrice jacobienne envisagée plus haut. On a donc, indépendamment du choix de toute base

$$\operatorname{div}\vec{V}(M) = \operatorname{trace}\left(d\vec{V}_M\right)$$

EXERCICE 23-1.25 Si $\vec{u}$ est un vecteur fixe et $\vec{V}$ est un champ de classe C^1, montrer que

$$\operatorname{div}\left(\vec{V}\wedge\vec{u}\right) = \vec{u}\,.\operatorname{rot}\vec{V}$$

De même, si $f : \mathcal{U} \to \mathbb{R}$ est un **champ scalaire** de classe C^2, son **laplacien** exprimé dans une base **orthonormale** de l'espace par

$$\Delta f(M) = \frac{\partial^2 f}{\partial x^2}(M) + \frac{\partial^2 f}{\partial y^2}(M) + \frac{\partial^2 f}{\partial z^2}(M)$$

ne dépend pas de la base puisque

$$\Delta f(M) = \operatorname{div}\left(\overrightarrow{\operatorname{grad}} f\right)(M)$$

REMARQUE 23-1.26 On suppose que l'espace $\mathbb{R}^n$ est muni de sa structure euclidienne canonique. On considère l'espace vectoriel $\mathbb{E} = C^\infty(\mathbb{R}^n, \mathbb{R})$ et pour $p \in \mathbb{N}^*$, on appelle opérateur différentiel linéaire homogène de degré p tout endomorphisme T de $\mathbb{E}$ de la forme

$$f \mapsto g = Tf \quad \text{avec} \quad \forall x \in \mathbb{R}^n \quad g(x) = \sum_{\substack{\alpha\in\mathbb{N}^n \\ |\alpha|=p}} a_\alpha(x)\,\frac{\partial^p f}{\partial x^\alpha}(x)$$

où les a_α sont des fonction C^∞ de $\mathbb{R}^n$ dans $\mathbb{R}$ (pour les notations, voir la section **18-2.5.2**). Soit T un tel opérateur (non identiquement nul) invariant par les translations et les automorphismes orthogonaux de $\mathbb{R}^n$, c'est à dire tel que

$$\forall f \in \mathbb{E} \quad T(f\circ\theta) = (Tf)\circ\theta$$

où θ est une translation ou une transformation orthogonale arbitraire de $\mathbb{R}^n$. On peut démontrer que ces hypothèses entraînent que p est nécessairement pair et que T est proportionnel à l'opérateur

$$\Delta^{\frac{p}{2}} = \Delta\circ\cdots\circ\Delta \quad \left(\frac{p}{2}\text{ fois}\right)$$

Δ étant l'opérateur laplacien défini par

$$\Delta f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

C'est pour cela que de nombreuses lois de la Physique, tenant compte de l'homogénéité et de l'isotropie de l'espace, ont une traduction mathématique faisant intervenir le laplacien.

23-1.4 Exemples d'applications

A titre d'exercice, nous donnons ici quelques résultats sur la théorie des fonctions dérivables d'une variable complexe. La définition a déjà été donnée à l'exercice **11-3**.41 : une fonction f définie sur un voisinage d'un complexe z_0 est dérivable en z_0 s'il existe $d \in \mathbb{C}$ tel que, pour $h \in \mathbb{C}$ avec $|h|$ assez petit on ait

$$f(z_0 + h) = f(z_0) + h\,d + |h|\,\varepsilon(h)$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$. Le complexe d est alors unique et on le note $f'(z_0)$. Nous savons par exemple que la somme d'une série entière est dérivable en tout point de son disque ouvert de convergence. Nous faisons le lien, dans cette section, avec les fonctions réelles de deux variables réelles :

EXERCICE 23-1.27 Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^2$ identifié à $\mathbb{C}$. On rappelle que, dans $\mathbb{R}^2$ euclidien canonique, l'endomorphisme canoniquement associé à la matrice

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

où $\mathbb{R}^2 \ni (a, b) \neq (0, 0)$ est une similitude directe (composée de l'homothétie de rapport $\sqrt{a^2 + b^2}$ et d'une rotation).

1. On se donne P et Q deux fonctions de classe C^1 définies sur $\mathcal{U}$ et à valeurs réelles

$$\mathcal{U} \ni (x, y) \mapsto P(x, y) \quad \text{et} \quad \mathcal{U} \ni (x, y) \mapsto Q(x, y)$$

On leur associe une fonction f de la variable complexe $z = x + iy$ en posant

$$f(x + iy) = P(x, y) + i\,Q(x, y)$$

Montrer que f est dérivable en tout point de $\mathcal{U}$ si et seulement si P et Q vérifient, dans $\mathcal{U}$, les **conditions de Cauchy-Riemann**

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \quad \text{et} \quad \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x}$$

Montrer qu'alors, en tout point de $\mathcal{U}$

$$f'(x + iy) = \frac{\partial P}{\partial x}(x, y) + i\,\frac{\partial Q}{\partial x}(x, y)$$

2. On suppose de plus les fonctions P et Q de classe C^2. Montrer que, si f est dérivable en tout point de $\mathcal{U}$, les fonctions P et Q sont **harmoniques**, c'est à dire que leur laplacien est identiquement nul dans $\mathcal{U}$:

$$\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} = \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} = 0$$

3. On suppose l'ouvert $\mathcal{U}$ étoilé et on se donne une fonction $A \in C^2(\mathcal{U}, \mathbb{R})$ harmonique. En considérant la forme différentielle définie sur $\mathcal{U}$ par

$$\omega = -\frac{\partial A}{\partial y} dx + \frac{\partial A}{\partial x} dy$$

montrer qu'on peut trouver une fonction $B \in C^2(\mathcal{U}, \mathbb{R})$ telle que la fonction F définie sur $\mathcal{U}$ par

$$F(x + i\,y) = A(x, y) + i\,B(x, y)$$

soit dérivable sur $\mathcal{U}$. Que peut-on dire de deux fonctions B_1 et B_2 répondant à la question ? Déterminer une fonction B dans le cas particulier

$$\mathcal{U} = \mathbb{C} - \mathbb{R}^- \quad \text{et} \quad A(x, y) = \frac{1}{2} \ln\left(x^2 + y^2\right)$$

et faire le lien avec l'exercice **9-6**.21.

4. Montrer que, si P et Q sont comme à la première question et si la fonction f' ne s'annule pas dans $\mathcal{U}$, la transformation de $\mathcal{U}$ dans $\mathbb{R}^2$ (euclidien canonique) définie par

$$\Phi : M = (x, y) \mapsto N = (P(x, y), Q(x, y))$$

transforme un arc γ de classe C^1 régulier tracé dans $\mathcal{U}$ en un arc régulier $\Phi(\gamma)$. Montrer que cette transformation est **conforme,** c'est à dire qu'elle conserve l'angle des tangentes à deux arcs réguliers sécants (voir l'exercice **18-2**.7). Que peut-on en déduire pour les familles de lignes de niveau[2] $(\Gamma_\alpha)_{\alpha \in \mathbb{R}}$ et $\left(\Gamma'_\beta\right)_{\beta \in \mathbb{R}}$ tracées dans $\mathcal{U}$ et définies par

$$\Gamma_\alpha = \{(x, y) \in \mathcal{U} \mid P(x, y) = \alpha\} \quad \text{et} \quad \Gamma'_\beta = \{(x, y) \in \mathcal{U} \mid Q(x, y) = \beta\}$$

[2]Il s'agit d'une situation que l'on rencontre en électrostatique, lorsqu'on étudie le potentiel créé par une répartition de charge invariante par translation dans une direction donnée choisie comme axe Oz (symétrie "cylindrique"). Le potentiel ne dépend pas de z et, en dehors des charges, vérifie

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$$

Si on peut trouver une fonction W conjuguée à V (c'est à dire telle que le couple (V, W) soit solution du problème résolu à la question 3), dans un plan perpendiculaire à Oz, les équipotentielles sont les lignes de niveaux de V tandis que les lignes de champ sont les lignes de niveau de W.

EXERCICE 23-1.28 (Suite de l'exercice précédent) : On suppose à présent que l'ouvert $\mathcal{U}$ est étoilé.

Si P et Q sont des fonctions de classe $\mathcal{C}^1$ vérifiant les conditions de Cauchy-Riemann, et si f est la fonction complexe dérivable associée, les formes différentielles

$$P\,dx - Q\,dy \quad \text{et} \quad Q\,dx + P\,dy$$

sont fermées et donc exactes dans $\mathcal{U}$. On introduit la notion de forme différentielle complexe[3]

$$\omega = (P\,dx - Q\,dy) + i\,(Q\,dx + P\,dy) = (P + i\,Q)\,(dx + idy)$$

que l'on représente symboliquement par

$$\omega = f(z)\,dz$$

On a donc : **si f est une fonction d'une variable complexe de classe $\mathcal{C}^1$ dans un ouvert $\mathcal{U}$ étoilé de $\mathbb{C}$, pour tout arc fermé γ continu et $\mathcal{C}^1$ par morceaux on a**

$$\int_\gamma f(z)\,dz = 0$$

Application : démonstration du théorème de D'Alembert.

Soit $P(z) = a_n\,z^n + a_{n-1}\,z^{n-1} + \cdots + a_0$ un polynôme à coefficients complexes, de degré n supposé ne pas posséder de racine complexe. On considère la fonction f définie par

$$f(z) = \frac{z^{n-1}}{P(z)}$$

Montrer que f est de classe $\mathcal{C}^1$ sur $\mathbb{C}$, et montrer que

$$\lim_{R\to+\infty} \int_{\gamma_R} f(z)\,dz = \frac{2i\pi}{a_n}$$

si γ_R est le chemin défini en affixe complexe par

$$\gamma_R : [0, 2\pi] \ni t \mapsto z_R(t) = R\,e^{it}$$

Conclure.

Le théorème étudié dans l'exercice qui suit est remarquable : il montre la grande différence entre la dérivabilité dans $\mathbb{R}$ et dans $\mathbb{C}$. Il s'agit également d'une belle illustration de l'efficacité des séries de Fourier :

[3] Qui n'est pas autre chose qu'un objet représenté formellement sous la forme

$$\omega = \omega_1 + i\omega_2$$

où ω_1 et ω_2 sont deux formes réelles. On intègre ω le long d'un arc paramétré γ par

$$\int_\gamma \omega = \int_\gamma \omega_1 + i\int_\gamma \omega_2$$

EXERCICE 23-1.29 Soit $\mathcal{U}$ un ouvert de $\mathbf{C}$ et $f : \mathcal{U} \to \mathbf{C}$ une fonction de classe C^1. Soit $z_0 \in \mathcal{U}$ et $R > 0$ tel que le disque fermé $D(z_0, R]$ de centre z_0 et de rayon R soit inclus dans $\mathcal{U}$. On se propose de montrer que, dans ce disque, f est somme d'une série entière de la variable $(z - z_0)$. En d'autres termes, nous aurons montré que toute fonction f de classe C^1 sur $\mathcal{U}$ est analytique dans $\mathcal{U}$, et que si

$$D(z_0, R] \subset \mathcal{U}$$

f est représentable dans tout ce disque par une série entière de la variable $(z - z_0)$.

1. Montrer que, pour $r \in [0, R]$, la fonction définie sur $\mathbf{R}$ par $\theta \mapsto f\left(z_0 + r\, e^{i\theta}\right)$ est somme de sa série de Fourier

$$\forall \theta \in \mathbf{R} \quad f\left(z_0 + r\, e^{i\theta}\right) = \sum_{n \in \mathbf{Z}} a_n(r)\, e^{in\theta}$$

et que cette série est normalement convergente sur $\mathbf{R}$.

2. En utilisant l'expression intégrale

$$a_n(r) = \frac{1}{2\pi} \int_0^{2\pi} f\left(z_0 + r\, e^{i\theta}\right) e^{-in\theta} d\theta$$

montrer que la fonction $r \mapsto a_n(r)$ est de classe C^1 sur $[0, R]$ et vérifie

$$\forall n \in \mathbf{Z} \quad \forall r \in [0, R] \quad r\, a_n'(r) = n\, a_n(r)$$

En déduire que $a_n(r) = 0$ pour $n < 0$ et est de la forme

$$a_n(r) = \alpha_n r^n$$

pour $n \geqslant 0$, avec $\alpha_n \in \mathbf{C}$. En déduire que, dans $D(z_0, R]$,

$$f(z) = \sum_{n=0}^{+\infty} \alpha_n (z - z_0)^n$$

Montrer que le rayon de convergence de cette série entière est supérieur ou égal à la distance de z_0 au complémentaire de $\mathcal{U}$.

REMARQUE 23-1.30 En utilisant le résultat de l'exercice **11-3**.41, on a montré que f est en fait C^∞ dans $\mathcal{U}$. On notera la grande différence avec la situation réelle. On démontrerait, avec beaucoup plus de difficulté, que les résultats qui précèdent sont encore valables si on suppose seulement f dérivable sur $\mathcal{U}$.

23-2 Intégrales doubles

Dans cette section, nous nous contenterons souvent d'énoncer les résultats sans démonstrations.

23-2.1 Compacts mesurables de $\mathbb{R}^2$

23-2.1.1 Mesure d'un compact pavable

On définit la mesure (ou aire) d'un pavé compact $R = [a, b] \times [c, d]$ de $\mathbb{R}^2$ par

$$m(R) = (b - a)(d - c)$$

On considère ensuite l'ensemble $\mathcal{E}$ des réunions de familles finies de tels pavés, auquel on ajoute l'ensemble vide. Il est clair qu'une union ou une intersection d'un nombre fini d'éléments de $\mathcal{E}$ est encore dans $\mathcal{E}$.

Un élément de $\mathcal{E}$ sera appelé "compact pavable" de $\mathbb{R}^2$. On démontre qu'un tel ensemble $E \neq \emptyset$ peut toujours s'écrire sous la forme

$$E = \bigcup_{i \in I} R_i$$

avec I finie, les R_i étant des pavés compacts "presque disjoints" c'est à dire tels que

$$i \neq j \Rightarrow R_i \cap R_j = \mathrm{Fr}(R_i) \cap \mathrm{Fr}(R_j)$$

où $\mathrm{Fr}(A)$ désigne la frontière d'une partie A de $\mathbb{R}^2$. Le réel

$$m(E) = \sum_{i \in I} m(R_i)$$

ne dépend alors pas de la décomposition de E, on l'appelle mesure (ou aire) du compact pavable E. Par convention, on pose $m(\emptyset) = 0$. On démontre :

PROPOSITION 23-2.1 L'application $m : \mathcal{E} \to \mathbb{R}^+$ **vérifie :**

$$\forall E, E' \in \mathcal{E} \quad E \subset E' \Rightarrow m(E) \leqslant m(E')$$
$$\forall E, E' \in \mathcal{E} \quad m(E \cup E') + m(E \cap E') = m(E) + m(E')$$

23-2.1.2 Compact mesurable

Soit K un compact de $\mathbb{R}^2$. Les ensembles

$$\mathcal{E}_K^+ = \{E \in \mathcal{E} \mid K \subset E\} \quad \text{et} \quad \mathcal{E}_K^- = \{E \in \mathcal{E} \mid E \subset K\}$$

sont non vides : K étant compact est inclus dans un pavé compact, qui appartient donc à $\mathcal{E}_K^+$, et $\mathcal{E}_K^-$ contient l'ensemble vide et toutes les parties finies de K. De plus, pour $E_1 \in \mathcal{E}_K^-$ et $E_2 \in \mathcal{E}_K^+$

$$E_1 \subset K \subset E_2 \Rightarrow m(E_1) \leqslant m(E_2)$$

Il en résulte que les ensembles $\{m(E) \mid E \in \mathcal{E}_K^-\}$ et $\{m(E) \mid E \in \mathcal{E}_K^+\}$ sont respectivement majorés et minorés et on a :

DÉFINITION 23-2.2 *Soit K un compact de $\mathbb{R}^2$. On appelle mesure intérieure de K le réel*

$$m_*(K) = \sup\{m(E) \mid E \in \mathcal{E}_K^-\}$$

et mesure extérieure de K le réel

$$m^*(K) = \inf\{m(E) \mid E \in \mathcal{E}_K^+\}$$

On a toujours

$$m_*(K) \leqslant m^*(K)$$

Lorsque $E \in \mathcal{E}$, on a clairement $m(E) = m_*(E) = m^*(E)$. On va donc prolonger la définition de m par :

DÉFINITION 23-2.3 *Un compact K du plan est mesurable (au sens de Jordan, on dit aussi quarrable) si et seulement si*

$$m_*(K) = m^*(K)$$

Ce réel positif est alors la mesure (ou l'aire) de K, on le note simplement $m(K)$.

La mesure ainsi définie est invariante par translation :

PROPOSITION 23-2.4 Si K est un compact mesurable et si $a \in \mathbb{R}^2$ est quelconque, le translaté $a + K$ est mesurable et

$$m(a + K) = m(K)$$

EXERCICE 23-2.5 On considère deux fonctions réelles f et g continues sur un segment $[a,b]$ de $\mathbb{R}$ vérifiant

$$\forall x \in [a,b] \quad f(x) \leqslant g(x)$$

et

$$K = \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \text{ et } f(x) \leqslant y \leqslant g(x)\}$$

Montrer que K est un compact mesurable et que

$$m(K) = \int_a^b [g(t) - f(t)]\, dt$$

(utiliser une subdivision à pas constant suffisamment petit de $[a,b]$ et des sommes de Riemann bien choisies pour f et g).

23-2.1.3 Propriétés

DÉFINITION 23-2.6 *Un compact K de $\mathbb{R}^2$ est négligeable si et seulement si $m^*(K) = 0$. Ceci équivaut à dire que, pour $\varepsilon > 0$ arbitraire, K peut être inclus dans une union finie de rectangles dont la somme des aires est inférieure à ε. Un tel compact est évidemment mesurable et de mesure nulle.*

Il est clair qu'une union d'un nombre fini de compacts négligeables est négligeable.

Les résultats qui suivent ont des démonstrations très techniques. Nous les admettrons :

PROPOSITION 23-2.7 Le support d'un arc paramétré de classe C^1

$$\gamma : [a,b] \ni t \mapsto M(t) \in \mathbb{R}^2$$

est négligeable.

PROPOSITION 23-2.8 Un compact K est mesurable si et seulement si sa frontière est négligeable.

Il en résulte qu'un compact dont la frontière est le support d'un arc continu et C^1 par morceaux sur un segment $[a,b]$ est mesurable.

PROPOSITION 23-2.9 Si K_1 et K_2 sont deux compacts mesurables, il en est de même de $K_1 \cup K_2$ et $K_1 \cap K_2$ et

$$m(K_1 \cup K_2) + m(K_1 \cap K_2) = m(K_1) + m(K_2)$$

L'ensemble des compacts mesurables de $\mathbb{R}^2$ est donc stable par unions et intersections finies.

PROPOSITION 23-2.10 Si $(K_i)_{i\in I}$ est une famille finie de compacts mesurables tels que

$$i \neq j \Rightarrow \overset{\circ}{K_i} \cap \overset{\circ}{K_j} = \emptyset$$

alors leur réunion est mesurable et

$$m\left(\bigcup_{i\in I} K_i\right) = \sum_{i\in I} m(K_i)$$

PROPOSITION 23-2.11 Soit K un compact mesurable de $\mathbb{R}^2$ et $u \in \mathcal{GL}(\mathbb{R}^2)$. Le compact $u(K)$ est mesurable et

$$m(u(K)) = |\det u|\ m(K)$$

En particulier, pour $\lambda \in \mathbb{R}$

$$m(\lambda K) = \lambda^2 m(K)$$

On pourra revoir à ce sujet la section **14-4.2** et la remarque **4-4.31**. Tout automorphisme de déterminant égal à ± 1 conserve donc la mesure dans $\mathbb{R}^2$. C'est le cas notamment pour les automorphismes orthogonaux de $\mathbb{R}^2$ euclidien canonique. On pourrait utiliser cette propriété, ainsi que l'invariance de la mesure par translation, pour définir la notion de compact mesurable dans un espace affine euclidien de dimension 2.

23-2.2 Intégrale d'une fonction continue

23-2.2.1 Intégrale d'une fonction étagée

DÉFINITION 23-2.12 *Si K est un compact mesurable, on appelle subdivision de K en compacts mesurables toute famille finie $d = (K_i)_{i\in I}$ de compacts mesurables tels que*

$$K = \bigcup_{i\in I} K_i \quad et \quad i \neq j \Rightarrow \overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset$$

Le pas de cette subdivision est, par définition,

$$\pi(d) = \max_{i\in I} \delta(K_i)$$

où $\delta(K_i)$ représente le diamètre de K_i.

DÉFINITION 23-2.13 *Soit K un compact mesurable de $\mathbb{R}^2$, une application $s : K \to \mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est étagée si et seulement s'il existe une subdivision $d = (K_i)_{i\in I}$ en compacts mesurables telle que s soit constante (égale à $\alpha_i \in \mathbb{K}$) sur chacun des intérieurs[4] des K_i. On définit alors l'intégrale de s sur K par*

$$\int_K s = \sum_{i\in I} \alpha_i m(K_i)$$

On peut voir que cette définition ne dépend pas de la subdivision adaptée à s : ceci se prouve en remarquant que, si $(K_i)_{i\in I}$ et $(K'_j)_{j\in J}$ sont deux telles subdivisions, la famille $(K_i \cap K'_j)_{(i,j)\in I\times J}$ est aussi adaptée à s, avec

$$\forall i \in I \quad m(K_i) = \sum_{j\in J} m(K_i \cap K'_j) \quad \text{et} \quad \forall j \in J \quad m(K'_j) = \sum_{i\in I} m(K_i \cap K'_j)$$

En utilisant la même méthode, on peut montrer que l'ensemble des fonctions étagées sur K à valeurs dans $\mathbb{K}$ est un sous-espace vectoriel de $\mathbb{K}^K$, et que l'application

$$s \mapsto I(s) = \int_K s$$

est une forme linéaire sur cet espace vectoriel, continue pour la norme de la convergence uniforme. En effet, par l'inégalité triangulaire, nous avons

$$\left| \int_K s \right| \leqslant \|s\|_\infty m(K)$$

[4]Si l'intérieur de K_i est vide, α_i n'est pas défini, mais il n'interviendra pas de manière effective dans la définition de l'intégrale de s, puisque dans ce cas

$$K_i = \mathrm{Fr}(K_i)$$

et donc $m(K_i) = 0$.

Si $m(K) = 0$, on pose simplement

$$\int_K s = 0$$

Si nous notons 1_K la fonction constante égale à 1 sur K, nous aurons évidemment

$$\int_K 1_K = m(K)$$

23-2.2.2 Intégrale d'une fonction continue

Nous définissons à présent l'intégrale d'une fonction continue

$$f : K \to \mathbb{K} = \mathbb{R} \text{ ou } \mathbb{C} \quad (x, y) \mapsto f(x, y)$$

définie sur un compact mesurable K de $\mathbb{R}^2$.

PROPOSITION 23-2.14 Si f est une fonction continue définie sur un compact mesurable K et à valeurs dans $\mathbb{K}$, il existe une suite $(s_n)_{n\in\mathbb{N}}$ de fonctions étagées sur K qui converge uniformément vers f sur K. La suite des intégrales

$$\left(\int_K s_n\right)_{n\in\mathbb{N}}$$

est alors convergente dans $\mathbb{K}$ et sa limite ne dépend pas de la suite $(s_n)_{n\in\mathbb{N}}$ considérée. On l'appelle intégrale de f sur K, et on note

$$\int_K f = \lim_{n\to+\infty} \int_K s_n$$

Démonstration : La convergence des intégrales (et le fait que la limite ne dépende pas de la suite choisie) utilise le critère de Cauchy et la majoration

$$\left|\int_K s_n - \int_K s_m\right| \leqslant \|s_n - s_m\|_\infty m(K)$$

L'existence de la suite $(s_n)_{n\in\mathbb{N}}$ est conséquence de la continuité uniforme de f sur K :

Si $\varepsilon > 0$ est fixé, on peut trouver un $\alpha > 0$ tel que

$$\forall m, m' \in K \quad \|m - m'\| \leqslant \alpha \Rightarrow |f(m) - f(m')| \leqslant \varepsilon$$

(on utilise par exemple la norme $\| \| = \| \|_\infty$ sur $\mathbb{R}^2$). On considère alors un pavé $[a, b] \times [c, d]$ tel que

$$K \subset [a, b] \times [c, d]$$

et on choisit $n \in \mathbb{N}^*$ pour avoir

$$\max\left(\frac{b-a}{n}, \frac{d-c}{n}\right) \leqslant \alpha$$

On écrit alors

$$[a, b] \times [c, d] = \bigcup_{(i,j)\in[0,n-1]^2} E_{i,j}$$

avec

$$E_{i,j} = \left[a + i\frac{b-a}{n}, a + (i+1)\frac{b-a}{n}\right] \times \left[c + j\frac{d-c}{n}, c + (j+1)\frac{d-c}{n}\right]$$

On a alors

$$K = \bigcup_{(i,j)\in[0,n-1]^2} K_{i,j}$$

où $K_{ij} = (E_{i,j} \cap K)$ est un compact mesurable (on ne tient compte effectivement que des K_{ij} non vides). On choisit un point $m_{i,j}$ dans $K_{i,j}$ lorsque cet ensemble est non vide, et on définit une fonction $s : K \to \mathbb{K}$ par $s(x) = f(x)$ si x appartient à la frontière d'un des $K_{i,j}$ et $s(x) = f(m_{i,j})$ si x est intérieur à un des $K_{i,j}$ (forcément unique). La fonction s est alors étagée sur K et vérifie clairement $\|f - s\|_\infty \leqslant \varepsilon$. ■

REMARQUE 23-2.15 On pourrait aussi intégrer de fonctions continues à valeurs dans un espace normé complet.

23-2.2.3 Propriétés

La démarche est la même que celle de la section **10-1.2.2** : on obtient toutes les propriétés de l'intégrale par passage à la limite :

- **L'intégrale est une forme linéaire sur $C^0(K, \mathbb{K})$, continue pour la norme de la convergence uniforme**

$$\forall f \in C^0(K, \mathbb{K}) \quad \left|\int_K f\right| \leqslant \|f\|_\infty \, m(K)$$

- **Intégrale d'une fonction complexe :**

$$\forall f \in C^0(K, \mathbb{C}) \quad \int_K f = \int_K \mathrm{Re}\, f + i \int_K \mathrm{Im}\, f$$

- **Forme linéaire positive :**

$$\forall f \in C^0(K, \mathbb{R}^+) \quad \int_K f \geqslant 0$$

De plus, s'il existe un point m_0 intérieur à K avec $f(m_0) > 0$, alors l'intégrale de f est strictement positive.

- **Inégalité du module :**

$$\forall f \in C^0(K, \mathbb{C}) \quad \left|\int_K f\right| \leqslant \int_K |f|$$

- **Intégration des inégalités :**

$$\forall f, g \in \mathcal{C}^0(K, \mathbb{R}) \quad f \leqslant g \text{ sur } K \Rightarrow \int_K f \leqslant \int_K g$$

On démontre, comme dans le cas de l'intégrale sur un segment, les inégalités de Schwarz et Minkowski :

$$\forall f, g \in \mathcal{C}^0(K, \mathbb{C}) \quad \begin{cases} \left|\int_K fg\right| \leqslant \left(\int_K |f|^2\right)^{\frac{1}{2}} \left(\int_K |g|^2\right)^{\frac{1}{2}} \\ \left(\int_K |f+g|^2\right)^{\frac{1}{2}} \leqslant \left(\int_K |f|^2\right)^{\frac{1}{2}} + \left(\int_K |g|^2\right)^{\frac{1}{2}} \end{cases}$$

- **Additivité par rapport à l'ensemble d'intégration :**

Si K_1 et K_2 sont deux compacts mesurables avec

$$\mathring{K}_1 \cap \mathring{K}_2 = \emptyset$$

et si $f : K_1 \cup K_2 \to \mathbb{C}$ est continue, on a

$$\int_{K_1 \cup K_2} f = \int_{K_1} f + \int_{K_2} f$$

- **Changement de variable linéaire :**

Soit K un compact mesurable de $\mathbb{R}^2$ et u un automorphisme de $\mathbb{R}^2$. Pour f continue de $u(K)$ dans $\mathbb{R}$ ou $\mathbb{C}$, on a

$$\int_{u(K)} f = |\det u| \int_K f \circ u$$

On montre cette égalité pour s étagée, le cas d'une fonction continue s'en déduisant par passage à la limite : si s est étagée sur $K' = u(K)$, considérons une subdivision de K' adaptée à s

$$K' = \bigcup_{i \in I} K'_i$$

On a, si α_i est la valeur de s sur l'intérieur de K'_i

$$\int_{u(K)} s = \sum_{i \in I} \alpha_i \, m(K'_i)$$

Si on note K_i le compact mesurable $u^{-1}(K'_i)$, on a

$$K = u^{-1}(K') = \bigcup_{i \in I} K_i$$

Comme u est un automorphisme, c'est en particulier un homéomorphisme qui transporte les ouverts. On a donc

$$i \neq j \Rightarrow \mathring{K}_i \cap \mathring{K}_j = u^{-1}\left(\mathring{K}'_i\right) \cap u^{-1}\left(\mathring{K}'_j\right) = u^{-1}\left(\mathring{K}'_i \cap \mathring{K}'_j\right) = \emptyset$$

Comme $m(K_i') = |\det u|\ m(K_i)$, on a

$$\int_{u(K)} s = \sum_{i\in I} \alpha_i\, m(K_i') = |\det u| \sum_{i\in I} \alpha_i m(K_i) = |\det u| \int_K s\circ u$$

puisque $(K_i)_{i\in I}$ est une subdivision de K adaptée à $s\circ u$ et α_i est la valeur prise par $s\circ u$ sur $\overset{\circ}{K_i}$.

23-2.3 Calcul des intégrales doubles

Nous donnons ici, sans démonstrations complètes, quelques techniques de calcul d'intégrales doubles.

23-2.3.1 Formule de Fubini

THÉORÈME 23-2.16 Soient deux fonctions réelles φ_1 et φ_2 continues sur un segment $[a,b]$ telles que

$$\forall x\in[a,b] \quad \varphi_1(x) \leqslant \varphi_2(x)$$

Le sous-ensemble K de $\mathbb{R}^2$ défini par

$$K = \left\{(x,y)\in\mathbb{R}^2 \mid x\in[a,b] \quad \textbf{et} \quad \varphi_1(x)\leqslant y\leqslant\varphi_2(x)\right\}$$

est un compact mesurable et, si

$$f : K\to\mathbb{C} \quad (x,y)\mapsto f(x,y)$$

est une fonction continue, on a

$$\int_K f = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\right) dx$$

Démonstration (partielle) : La mesurabilité de K a été évoquée à l'exercice **23-2**.5. Il est facile de voir qu'avec les hypothèses faites sur φ_1, φ_2 et f, l'application

$$[a,b]\ni x\mapsto \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy$$

est continue sur $[a,b]$, ce qui justifie l'existennce de l'intégrale

$$\int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\right) dx$$

Nous ferons la démonstration dans le cas particulier où

$$K = [a,b]\times[c,d]$$

c'est à dire lorsque les fonctions φ_1 et φ_2 sont constantes. On fera ainsi le lien avec le résultat du théorème **10-2**.7. Notons $K_x = [a,x] \times [c,d]$ et

$$\forall x \in [a,b] \quad g(x) = \int_{K_x} f$$

(abus d'écriture commode, on considère en fait la restriction de la fonction f à K_x). Par additivité de l'intégrale par rapport à l'ensemble d'intégration, on a clairement

$$\forall x, x' \in [a,b] \quad |g(x) - g(x')| \leqslant |x - x'|\,(d-c)\,\|f\|_\infty$$

La fonction g est donc continue sur $[a,b]$. Nous allons prouver que g est de classe C^1 sur $[a,b]$, de dérivée

$$g'(x) = \int_c^d f(x,y)\,dy$$

et l'égalité

$$g(b) = g(b) - g(a) = \int_a^b g'(x)\,dx$$

donnera le résultat. Supposons donc x et $x+h \in [a,b]$, avec par exemple $h > 0$ (le cas $h < 0$ se traiterait de la même manière), et étudions

$$g(x+h) - g(x) = \int_{[x,x+h]\times[c,d]} f$$

Notons f_1 l'application de $[a,b] \times [c,d]$ dans $\mathbb{C}$ définie par

$$\forall (t,y) \in K \quad f_1(t,y) = f(x,y)$$

Cette fonction est bien continue et, par continuité uniforme de f, pour $\varepsilon > 0$ arbitrairement choisi, on peut trouver un $\alpha > 0$ tel que

$$0 < h \leqslant \alpha \Rightarrow [\forall (t,y) \in [x,x+h] \times [c,d] \quad |f(t,y) - f_1(t,y)| \leqslant \varepsilon]$$

On en déduit, par inégalité du module,

$$0 < h \leqslant \alpha \Rightarrow \left| \int_{[x,x+h]\times[c,d]} f - \int_{[x,x+h]\times[c,d]} f_1 \right| \leqslant \varepsilon\,(d-c)\,h$$

On a donc obtenu

$$g(x+h) - g(x) \underset{h\to 0}{=} \int_{[x,x+h]\times[c,d]} f_1 + o(h)$$

Pour approcher l'intégrale de la fonction f_1 (qui ne dépend que de la variable y), on considère la décomposition

$$[x,x+h] \times [c,d] = \bigcup_{(i,j)\in[0,n-1]^2} E_{i,j}$$

avec

$$E_{i,j} = \left[x + i\frac{h}{n}, x + (i+1)\frac{h}{n}\right] \times \left[c + j\frac{d-c}{n}, c + (j+1)\frac{d-c}{n}\right]$$

et une fonction étagée s_n coïncidant avec f_1 sur chacune des frontières de ces rectangles, et valant

$$\alpha_{i,j} = f_1\left(x + i\frac{h}{n}, c + j\frac{d-c}{n}\right) = f\left(x, c + j\frac{d-c}{n}\right)$$

sur l'intérieur de $E_{i,j}$. La suite $(s_n)_{n\in\mathbb{N}^*}$ converge uniformément vers f_1 sur $[x, x+h] \times [c,d]$, avec

$$\int_{[x,x+h]\times[c,d]} s_n =$$
$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \frac{h}{n}\frac{d-c}{n} f\left(x, c + j\frac{d-c}{n}\right) = h\sum_{j=0}^{n-1}\frac{d-c}{n} f\left(x, c + j\frac{d-c}{n}\right)$$

On a donc

$$\lim_{n\to+\infty} \int_{[x,x+h]\times[c,d]} s_n = h\int_c^d f(x,y)\,dy = \int_{[x,x+h]\times[c,d]} f_1$$

(propriétés des sommes de Riemann des fonctions continues), et finalement

$$g(x+h) - g(x) = h\int_c^d f(x,y)\,dy + o(h)$$

ce qui permet de conclure aisément. ■

REMARQUE 23-2.17 Le raisonnement qui précède est une autre démonstration du théorème **10-2**.7, puisqu'on peut évidemment intervertir les rôles joués par x et y. Dans le cas général, si K admet une écriture de la forme

$$K = \left\{(x,y) \in \mathbb{R}^2 \mid y \in [c,d] \quad \text{et} \quad \psi_1(y) \leqslant x \leqslant \psi_2(y)\right\}$$

avec $\psi_1 \leqslant \psi_2$ deux fonctions continues sur $[c,d]$, on aura aussi

$$\int_K f = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx\right) dy$$

Lorsque K se décompose en une union finie de compacts d'intérieurs disjoints deux à deux et admettant chacun une telle description, on peut opérer par additivité par rapport à l'ensemble d'intégration.

Notation : l'intégrale d'une fonction f continue sur un compact mesurable est aussi notée

$$\iint_K f(x,y)\,dxdy$$

où, comme dans le cas de l'intégrale en dimension 1, les lettres x et y sont "muettes".

EXERCICE 23-2.18 Calculer l'intégrale

$$I = \iint_K \frac{2xy}{1+x^2+y^2}\,dxdy$$

où K est le compact du plan

$$K = \left\{(x,y) \in \mathbb{R}^2 \mid y \geqslant 0 \quad \text{et} \quad x^2+y^2-x \leqslant 0\right\}$$

(Réponse : $I = \frac{3}{4} - \ln 2$).

Dans la pratique, on est souvent limité dans l'utilisation de la formule de Fubini par la complexité du domaine K. On utilisera souvent un changement de variable pour transformer K en un domaine "plus simple" : par exemple, si

$$K = \left\{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0,\, y \geqslant 0 \quad \text{et} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}$$

il serait très maladroit d'appliquer la formule de Fubini pour calculer

$$I = \iint_K \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^3 dxdy$$

en écrivant

$$I = \int_0^a \left(\int_0^{b\sqrt{1-\frac{x^2}{a^2}}} \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^3 dy\right) dx$$

Un changement de variable très simple va permettre de calculer cette intégrale à moindre frais :

23-2.3.2 Changement de variable

Nous admettrons le résultat suivant :

PROPOSITION 23-2.19 Soient K un compact mesurable, U un voisinage ouvert de K et $\Phi : U \to \mathbb{R}^2$ une application de classe C^1, dont la restriction à l'intérieur de K induit un C^1 difféomorphisme[5]

$$\Phi|_{\overset{\circ}{K}} : \overset{\circ}{K} \to \Phi\left(\overset{\circ}{K}\right)$$

Le compact $\Phi(K)$ est alors mesurable et, si $f : \Phi(K) \to \mathbb{C}$ est continue, on a

$$\int_{\Phi(K)} f = \int_K (f \circ \Phi)\, |\det J_\Phi|$$

où J_Φ représente la matrice jacobienne de Φ.

[5]On sait que cela signifie que Φ est injective en restriction à $\overset{\circ}{K}$, et que sa différentielle en tout point de cet intérieur est un automorphisme de $\mathbb{R}^2$.

Si Φ est définie par

$$U \ni (u,v) \mapsto \Phi(u,v) = (x(u,v), y(u,v))$$

la formule de changement de variable pourra s'écrire aussi

$$\iint_{\Phi(K)} f(x,y)\,dxdy$$
$$= \iint_{K} f(x(u,v), y(u,v)) \left| \frac{\partial x}{\partial u}(u,v)\frac{\partial y}{\partial v}(u,v) - \frac{\partial x}{\partial v}(u,v)\frac{\partial y}{\partial u}(u,v) \right| dudv$$

Sous réserve de vérification des hypothèses énoncées précédemment, cette transformation se fait de manière "automatique", en remplaçant x et y par leurs expressions en fonction de u et v, et en remplaçant le symbole "$dxdy$" par "$|\det J_\Phi(u,v)|\ dudv$".

REMARQUE 23-2.20 Dans cette formule, la valeur absolue autour du déterminant jacobien peut surprendre : il n'y a en effet pas de valeur absolue dans la formule de changement de variable (en dimension 1) énoncée au théorème **10-1.50**. C'est parce que dans ce cas, les intégrales sont calculées sur des intervalles orientés : on a donné un sens à

$$\int_{\beta}^{\alpha} f(\varphi(t))\,\varphi'(t)\,dt$$

lorsque $\beta > \alpha$. Si on rétablit les bornes lorsque φ décroît, on change φ' en $-\varphi'$.

Lorsque $\overset{\circ}{K}$ est connexe, la fonction (continue) det J_Φ ne s'annulant pas sur $\overset{\circ}{K}$, elle garde un signe constant. La valeur absolue autour du déterminant rétablit alors le signe +, de manière à ce que l'intégrale

$$\int_K (f \circ \Phi)\ |\det J_\Phi|$$

soit positive si f l'est sur $\Phi(K)$.

Sans donner de démonstration de cette "formule", nous pouvons cependant comprendre d'où elle vient : si $(K_i')_{i\in I}$ est une subdivision de $\Phi(K)$ en compacts mesurables de diamètres "petits" et si on choisit pour tout i un point $m_i \in K_i'$, une fonction étagée valant $f(m_i)$ sur l'intérieur de K_i' est une "bonne" approximation uniforme de f (si cette fonction coïncide avec f sur les frontières des compacts de subdivision) et donc

$$\int_{\Phi(K)} f \approx \sum_{i\in I} f(m_i)\, m(K_i')$$

Si Φ induit un homéomorphisme de K sur $\Phi(K)$, $K_i = \Phi^{-1}(K_i')$ est un compact de K de diamètre petit ($\Phi|_K$ et $\Phi|_{\Phi(K)}^{-1}$ sont uniformément continues). Si on considère $p_i = \Phi^{-1}(m_i) \in K_i$, on peut approcher $\Phi|_{K_i}$ par son application affine tangente en p_i. On a donc

$$m(K_i') = m(\Phi(K_i)) \approx |\det J_\Phi(p_i)|\ m(K_i)$$

On obtiendra

$$\int_{\Phi(K)} f \approx \sum_{i\in I} f\left(\Phi\left(p_i\right)\right) \left|\det J_\Phi\left(p_i\right)\right| m\left(K_i\right) \approx \int_K \left(f \circ \Phi\right) \ \left|\det J_\Phi\right|$$

EXEMPLE DU PASSAGE EN COORDONNEES POLAIRES :
Soit D un compact du plan admettant une représentation polaire

$$D = \{(r\ \cos\theta, r\ \sin\theta) \mid (r,\theta) \in K\}$$

où K est un compact mesurable de $\mathbb{R}^2$ vérifiant (par exemple)

$$K \subset \mathbb{R}^+ \times [0, 2\pi]$$

Les hypothèses du théorème sont alors vérifiées pour l'application (de classe C^∞)

$$\Phi : \mathbb{R}^2 = U \to \mathbb{R}^2 \quad \Phi\left(r,\theta\right) = \left(r\cos\theta, r\sin\theta\right)$$

puisque $\overset{\circ}{K} \subset \mathbb{R}^{*+} \times]0, 2\pi[$, ouvert sur lequel Φ induit un C^1 difféomorphisme. Comme le déterminant jacobien de Φ en (r,θ) vaut $r \geqslant 0$, nous aurons donc, pour f continue dans D :

$$\iint_D f\left(x,y\right)\,dxdy = \iint_K f\left(r\cos\theta, r\sin\theta\right)\,r\,drd\theta$$

EXERCICE 23-2.21 Reprendre le calcul de l'intégrale de l'exercice **23-2.18** en passant en coordonnées polaires. On trouvera

$$I = \iint_K \frac{2xy}{1+x^2+y^2}\,dxdy = \int_0^{\frac{\pi}{2}} \left(\int_0^{\cos\theta} \frac{r^3}{1+r^2}dr\right) 2\cos\theta\sin\theta\,d\theta$$

Calculer cette dernière intégrale.

EXERCICE 23-2.22 Si

$$K = \left\{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0,\ y \geqslant 0 \quad \text{et} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}$$

calculer

$$I = \iint_K \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^3 dxdy$$

EXERCICE 23-2.23 On définit l'application f sur $\mathbb{R}^2$ par

$$f\left(x,y\right) = e^{-\left(x^2+y^2\right)}$$

Pour $R > 0$, on pose

$$K_R = [0,R] \times [0,R] \text{ et } C_R = \left\{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0,\ y \geqslant 0 \text{ et } x^2+y^2 \leqslant R^2\right\}$$

Montrer que

$$\forall R > 0 \quad \int_{C_R} f \leqslant \int_{K_R} f \leqslant \int_{C_{R\sqrt{2}}} f$$

En déduire la valeur de

$$\int_0^{+\infty} e^{-t^2}\,dt$$

23-2.3.3 Formule de Green-Riemann

Soit $\gamma : [a,b] \ni t \mapsto M(t) \in \mathbb{R}^2$ un arc continu et C^1 par morceaux, fermé (c'est à dire $M(a) = M(b)$) et simple (l'application M restreinte à $[a,b[$ est injective). Son support $S(\gamma)$ est une partie compacte de $\mathbb{R}^2$. On démontre que son complémentaire dans $\mathbb{R}^2$ est réunion de deux ouverts connexes disjoints, dont l'un O_1 est borné. Le domaine

$$I(\gamma) = O_1 \cup S(\gamma)$$

est alors un compact mesurable du plan, appelé **intérieur** de γ. On admet qu'on peut définir une orientation de γ correspondant à la notion intuitive de description du support "dans le sens trigonométrique direct", et on suppose que le paramétrage choisi correspond à cette orientation (intuitivement, l'intérieur de γ est localement "à gauche" d'un observateur décrivant le support selon la loi $t \mapsto M(t)$). On dit alors que γ est orienté positivement.

PROPOSITION 23-2.24 Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^2$ et $\gamma : [a,b] \ni t \mapsto M(t) \in \mathbb{R}^2$ un arc continu et C^1 par morceaux, fermé et simple, orienté positivement, tel que

$$I(\gamma) \subset \mathcal{U}$$

Si P et Q sont des fonctions de classe C^1 de $\mathcal{U}$ dans $\mathbb{R}$, on a

$$\int_\gamma P(x,y)\,dx + Q(x,y)\,dy = \iint_{I(\gamma)} \left(\frac{\partial Q}{\partial x}(x,y) - \frac{\partial P}{\partial y}(x,y)\right) dxdy$$

On remarquera que l'hypothèse faite sur γ impose à l'intérieur de γ (et pas seulement le support) d'être dans $\mathcal{U}$. Cette hypothèse assure l'existence de l'intégrale double[6]. Nous admettrons ce résultat dans le cas général, en la justifiant dans un cas particulier :

> *Justification (partielle)* : Nous supposons que l'intérieur de γ peut être représenté par
>
> $$K = \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \quad \text{et} \quad \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$$
> $$\text{et} \quad K = \{(x,y) \in \mathbb{R}^2 \mid y \in [c,d] \quad \text{et} \quad \psi_1(y) \leqslant x \leqslant \psi_2(y)\}$$
>
> où $\varphi_1 \leqslant \varphi_2$ sont deux fonctions continues et C^1 par morceaux sur $[a,b]$, $\psi_1 \leqslant \psi_2$ vérifiant la même propriété sur $[c,d]$. L'arc γ utilisé est alors C^1 par morceaux, et est défini par la juxtaposition des chemins
>
> $$\begin{cases} \gamma_1 : [a,b] \ni t \mapsto M_1(t) = (t, \varphi_1(t)) \\ \gamma_2 : [\varphi_1(b), \varphi_2(b)] \ni t \mapsto M_2(t) = (b,t) \\ \gamma_3 : [a,b] \ni t \mapsto M_3(t) = (a+b-t, \varphi_2(a+b-t)) \\ \gamma_4 : [\varphi_1(a), \varphi_2(a)] \ni t \mapsto M_4(t) = (a, \varphi_2(a) + \varphi_1(a) - t) \end{cases}$$

[6] La formule de Green-Riemann est un cas particulier d'un résultat plus général (formule de Stokes) permettant de remplacer une intégrale "à l'intérieur" d'un domaine par une intégrale "au bord". La version la plus élémentaire de la formule de Stokes est tout simplement

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

pour f de classe C^1 sur un segment $[a,b]$.

Sans trop nous préoccuper de la validité du changement de paramétrage, nous aurons une représentation analogue de γ en permutant les rôles joués par x et y (en conservant l'orientation).

A l'aide de la formule de Fubini, nous évaluons

$$\iint_K \frac{\partial P}{\partial y}(x,y)\, dxdy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}(x,y)\, dy \right) dx$$

Comme P est supposée de classe $\mathcal{C}^1$ sur $\mathcal{U}$, en utilisant la définition de la dérivée partielle on a

$$\forall x \in [a,b] \quad \int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}(x,y)\, dy = P(x,\varphi_2(x)) - P(x,\varphi_1(x))$$

et donc

$$-\iint_K \frac{\partial P}{\partial y}(x,y)\, dxdy = \int_a^b P(t,\varphi_1(t))\, dt - \int_a^b P(t,\varphi_2(t))\, dt$$

On obtient finalement

$$-\iint_K \frac{\partial P}{\partial y}(x,y)\, dxdy = \int_{\gamma_1} P(x,y)\, dx + \int_{\gamma_3} P(x,y)\, dx = \int_\gamma P(x,y)\, dx$$

puisqu'on a clairement

$$\int_{\gamma_2} P(x,y)\, dx = \int_{\gamma_4} P(x,y)\, dx = 0$$

On montrerait de même, en permutant les rôles joués par x et y, que

$$\int_\gamma Q(x,y)\, dy = \iint_K \frac{\partial Q}{\partial x}(x,y)\, dxdy$$

En soustrayant ces inégalités, on obtient le résultat souhaité. ■

EXEMPLE 23-2.25 Pour calculer l'intégrale

$$I = \iint_K x^2\, dxdy$$

où K est le compact du plan défini par

$$K = \left\{ (x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1 \right\}$$

on pourra utiliser le paramétrage de l'ellipse

$$[0,2\pi] \ni t \mapsto M(t) = (a\cos t, b\sin t)$$

et choisir (par exemple) $Q(x,y) = \frac{x^3}{3}$ et $P(x,y) = 0$. On obtient alors

$$I = \frac{a^3 b}{3} \int_0^{2\pi} \cos^4 t \, dt = \frac{1}{4} a^3 b \pi$$

Un changement de variables ($x = ar\cos\theta$, $y = br\sin\theta$ avec $(r,\theta) \in [0,1] \times [0,2\pi]$) transformerait "$dxdy$" en "$abr\,drd\theta$" et donnerait

$$I = \iint_{[0,1]\times[0,2\pi]} a^3 b r^3 \cos^2\theta \, drd\theta = a^3 b \left(\int_0^1 r^3 dr\right)\left(\int_0^{2\pi} \cos^2\theta \, d\theta\right) = \frac{1}{4} a^3 b\pi$$

APPLICATION : CALCUL DE L'AIRE D'UN SECTEUR PLAN. Soit K un compact mesurable pouvant être décrit comme intérieur d'un arc fermé simple et C^1 par morceaux. En prenant pour fonctions $P(x,y) = -\frac{y}{2}$ et $Q(x,y) = \frac{x}{2}$, nous aurons

$$\frac{\partial Q}{\partial x}(x,y) - \frac{\partial P}{\partial y}(x,y) = 1$$

sur K, et donc

$$K = I(\gamma) \Rightarrow m(K) = \frac{1}{2} \int_\gamma x dy - y dx$$

EXEMPLE 23-2.26 Aire du domaine plan délimité par l'astroïde définie paramétriquement par

$$\forall t \in [0,2\pi] \quad x(t) = a\cos^3 t \quad \text{et} \quad y(t) = a\sin^3 t$$

(Réponse : $\frac{3}{8} a^2 \pi$).

Lorsque le secteur K est déterminé en coordonnées polaires par les conditions

$$\theta \in [\theta_1, \theta_2] \quad \text{et} \quad 0 \leqslant r \leqslant f(\theta)$$

où f est une fonction positive de classe C^1 par morceaux (en fait le résultat subsisterait avec f simplement continue) sur $[\theta_1, \theta_2]$, l'arc γ est obtenu comme juxtaposition de trois arcs. L'intégrale de la forme différentielle $xdy - ydx$ le long de OM_1 et de M_2O est nulle (paramétrage de la forme $x = \alpha t$ et $y = \beta t$, avec α et β constants). Le long de l'arc M_1M_2, on a la représentation paramétrique

$$x(\theta) = f(\theta)\cos\theta \quad \text{et} \quad y(\theta) = f(\theta)\sin\theta$$

et donc

$$xdy - ydx = f^2(\theta)\, d\theta$$

On en déduit

$$m(K) = \frac{1}{2} \int_{\theta_1}^{\theta_2} f^2(\theta)\, d\theta$$

EXEMPLE 23-2.27 Aire délimitée par la cardioïde d'équation polaire

$$r = a\,(1 + \cos\theta)$$

(Réponse : $\dfrac{3}{2}a^2\pi$).

EXERCICE 23-2.28 (Inégalité isopérimétrique) Dans $\mathbb{R}^2$ euclidien canonique, on considère un arc de classe $\mathcal{C}^1$ régulier fermé simple dont le support a pour longueur L. On paramètre cet arc par son abscisse curviligne

$$[0, L] \ni s \mapsto M\,(s) = (x\,(s)\,, y\,(s))$$

où x et y sont de classe C^1 et vérifient $x\,(0) = x\,(L)$, $x'\,(0) = x'\,(L)$ et $y\,(0) = y\,(L)$, $y'\,(0) = y'\,(L)$. On a de plus, puisque le paramétrage est normal,

$$\forall\, s \in [0, L] \quad x'^2\,(s) + y'^2\,(s) = 1$$

On pose

$$\forall\, s \in [0, L] \quad z\,(s) = x\,(s) + i\,y\,(s)$$

Montrer qu'on peut prolonger z par L-périodicité en une fonction de classe $\mathcal{C}^1$. Montrer que l'aire du domaine délimité par γ est donnée par

$$A = \frac{1}{2}\,\mathrm{Im}\left(\int_0^L z'\,(s)\,\overline{z}\,(s)\; ds\right)$$

En utilisant la suite $(c_n)_{n\in\mathbb{Z}}$ des coefficients de Fourier de z, montrer que

$$A = \pi \sum_{n\in\mathbb{Z}} n\; |c_n|^2$$

Montrer également que

$$\sum_{n\in\mathbb{Z}} n^2\,|c_n|^2 = \frac{L^2}{4\pi^2}$$

et en déduire l'inégalité

$$A \leqslant \frac{L^2}{4\pi}$$

Déterminer le support de γ dans le cas de l'égalité.

23-2.4 Extension aux intégrales triples et multiples

La notion de compact mesurable peut se définir dans $\mathbb{R}^n$, les ensembles pavables étant encore les unions finies de pavés compacts, produits cartésiens de n segments, avec la mesure définie par

$$m\left(\prod_{i=1}^{n} [a_i, b_i]\right) = \prod_{i=1}^{n} (b_i - a_i)$$

Les résultats obtenus dans $\mathbb{R}^2$ se généralisent, avec quelques modifications d'énoncés. Par exemple, la proposition **23-2.7** devient :

PROPOSITION 23-2.29 Si U est un ouvert de $\mathbb{R}^{n-1}$ et si $M : U \to \mathbb{R}^n$ est une application de classe $\mathcal{C}^1$, pour tout K compact inclus dans U le compact $M(K)$ est négligeable.

En particulier, dans $\mathbf{R}^3$, toute morceau compact de support de nappe paramétrée de classe $\mathcal{C}^1$ est négligeable.

On construit l'intégrale d'une fonction continue sur un compact mesurable de $\mathbb{R}^n$ et le théorème de Fubini s'écrit, dans le cas particulier $n = 3$:

PROPOSITION 23-2.30 (Intégration par piles) Soit K un compact mesurable de $\mathbb{R}^3$ et $f : K \to \mathbb{C}$ une fonction continue. On suppose que K admet une description de la forme

$$K = \left\{(x,y,z) \in \mathbf{R}^3 \mid (x,y) \in K_1 \text{ et } \varphi_1(x,y) \leqslant z \leqslant \varphi_2(x,y)\right\}$$

où K_1 est un compact mesurable de $\mathbf{R}^2$ et $\varphi_1 \leqslant \varphi_2$ sont deux applications continues de K_1 dans $\mathbb{R}$. On a alors

$$\int_K f = \iint_{K_1} \left(\int_{\varphi_1(x,y)}^{\varphi_2(x,y)} f(x,y,z)\, dz \right) dxdy$$

PROPOSITION 23-2.31 (Intégration par plaques) On suppose cette fois que K admet une description de la forme

$$K = \left\{(x,y,z) \in \mathbb{R}^3 \mid z \in [a,b] \quad \text{et} \quad (x,y) \in K_z\right\}$$

où, pour tout $z \in [a,b]$, K_z est un compact mesurable de $\mathbf{R}^2$. Si l'application

$$z \mapsto \iint_{K_z} f(x,y,z)\, dxdy$$

est continue par morceaux sur $[a,b]$, alors

$$\int_K f = \int_a^b \left(\iint_{K_z} f(x,y,z)\, dxdy \right) dz$$

On a évidemment des résultats analogues en faisant jouer un rôle particulier à la variable x ou à la variable y. Pour ces raisons, l'intégrale de f est aussi notée

$$\int_K f = \iiint_K f(x,y,z)\, dxdydz$$

EXERCICE 23-2.32 En utilisant les deux méthodes, calculer

$$\iiint_K x(y+z)\, dxdydz$$

où $K = \left\{(x,y,z) \in \mathbb{R}^3 \mid x \geqslant 0,\ y \geqslant 0,\ z \geqslant 0 \quad \text{et} \quad x+y+z \leqslant 1\right\}$. (Réponse : $\frac{1}{60}$).

La formule du changement de variables s'énonce comme dans la cas de $\mathbb{R}^2$. Pour $n = 3$, on a en particulier :

- **Passage en coordonnées cylindriques d'axe Oz :**
Si K est un compact mesurable de $\mathbb{R}^3$ admettant la description

$$K = \left\{(r\cos\theta, r\sin\theta, z) \in \mathbb{R}^3 \mid (r,\theta,z) \in K_1\right\}$$

où K_1 est un compact mesurable inclus dans $\mathbb{R}^+ \times [0, 2\pi] \times \mathbb{R}$, on a

$$\iiint_K f(x,y,z)\; dxdydz = \iiint_{K_1} f(r\cos\theta, r\sin\theta, z)\; r\, drd\theta dz$$

- **Passage en coordonnées sphériques d'axe Oz :**
Si K est un compact mesurable de $\mathbb{R}^3$ admettant la description

$$K = \left\{(r\cos\theta\cos\varphi, r\sin\theta\cos\varphi, r\sin\varphi) \in \mathbb{R}^3 \mid (r,\theta,\varphi) \in K_1\right\}$$

où K_1 est un compact mesurable inclus dans $\mathbb{R}^+ \times [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, on a

$$\iiint_K f(x,y,z)\; dxdydz = \iiint_{K_1} f(r\cos\theta\cos\varphi, r\sin\theta\cos\varphi, r\sin\varphi)\; r^2\cos\varphi\, drd\theta d\varphi$$

EXERCICE 23-2.33 Si $a > 0$ est un réel, calculer

$$\iiint_K xy\, dxdydz$$

où le domaine K est défini par

$$K = \left\{(x,y,z) \in \mathbb{R}^3 \mid y \geqslant 0,\ z \leqslant x + y \quad \text{et} \quad x^2 + y^2 - 2ax \leqslant 0\right\}$$

(Réponse : $\left(\frac{4}{5} + \frac{\pi}{8}\right)a^5$).

EXERCICE 23-2.34 Même question avec

$$K = \left\{(x,y,z) \in \mathbb{R}^3 \mid y \geqslant 0,\ z \geqslant 0 \quad \text{et} \quad x^2 + y^2 + z^2 - 2ax \leqslant 0\right\}$$

(Réponse : $\frac{\pi}{8}a^5$).

23-3 Exercices

EXERCICE 23-3.1 Soit $\mathcal{D} = \{(x,y) \in \mathbb{R}^2 \mid \frac{1}{\sqrt{2}}|x-y| \leqslant x^2 + y^2 \leqslant 1\}$ Calculer

$$\iint_{\mathcal{D}} (x^3 + y^3) dxdy \text{ et } \iint_{\mathcal{D}} (x^2 + y^2) dxdy$$

EXERCICE 23-3.2 Calculer $\iint_D \sqrt{xy}\, dxdy$ où

$$D = \{(x,y) \in \mathbb{R}^2 \mid y \geqslant 0,\ (x+y)^2 \leqslant 2x\}$$

EXERCICE 23-3.3 Soit $F : [0,1] \times [0,1] \to \mathbf{R}$ continue. Déterminer

$$\lim_{n \to +\infty} \iint_{[0,1]^2} F(x,y)\cos(nxy)dxdy$$

EXERCICE 23-3.4 Calculer $\iiint_D \frac{dx\,dy\,dz}{a^2+x^2+2y^2+3z^2}$ où

$$D = \{(x,y,z) \in \mathbf{R}^3 \mid x^2+2y^2+3z^2 \leqslant a^2\}$$

EXERCICE 23-3.5 Calcul d'intégrales doubles : calculer $\iint_D f(x,y)dxdy$ dans les cas suivants :

- $f(x,y) = 1/\sqrt{y^2-x^2}$ et $D = \{(x,y)/\, x > y,\ y^2 < x\}$
- $f(x,y) = x^2+y^2$ et $D = \{(x,y)/\, 0 \leqslant x \leqslant 2a,\ 0 \leqslant y \leqslant \sqrt{2ax-x^2}\}$ $(a>0)$
- $f(x,y) = \sqrt{\frac{1-x^2-y^2}{1+x^2+y^2}}$ et $D = \{(x,y)/\, x \geqslant 0,\ y \geqslant 0,\ x^2+y^2 \leqslant R^2\}$ $(R>0)$
- $f(x,y) = 1/(1+x^2+y^2)^2$ et $D = \{(x,y)/\, |x| \leqslant x^2+y^2 \leqslant 1\}$
- $f(x,y) = x^2y\sqrt[3]{a^3-x^3-y^3}$ et $D = \{(x,y)/\, x \geqslant 0,\ y \geqslant 0,\ x^3+y^3 \leqslant a^3\}$.

EXERCICE 23-3.6 Calcul d'aires de domaines plans :

- $D = \{(x,y)/\, x \geqslant 0,\ y \geqslant 0,\ xy \geqslant 1/4,\ x^2+y^2 \leqslant 1\}$
- $D = \{(x,y)/\, \frac{x^2}{a^2}+\frac{y^2}{b^2} \leqslant 1,\ \frac{y^2}{a^2}+\frac{x^2}{b^2} \leqslant 1\}$

EXERCICE 23-3.7 Calcul d'intégrales triples : calculer $\iiint_D f(x,y,z)dxdydz$ dans les cas suivants :

- $f(x,y,z) = x^2+y^2$ et $D = \{(x,y,z)/\, z \geqslant 0,\ a^2 \leqslant x^2+y^2+z^2 \leqslant b^2\}$
- $f(x,y,z) = (ax+by+cz)^2$ et $D = \{(x,y,z)/\, x^2+y^2+z^2 \leqslant R^2\}$

EXERCICE 23-3.8 Intégrales de Fresnel :

1. Soit f continue de $[0,\pi/2]$ dans $\mathbf{R}^{+*}$. Pour tout $t>0$, on pose :

   $D_t = \{(x,y) \in \mathbb{R}^2/\, \exists(\theta,r) \in [0,\pi/2]\times[0,tf(\theta)],$ tel que $x = r\cos\theta,\ y = r\sin\theta\}$

   On appelle $\phi(t) = \iint_{D_t} \sin(x^2+y^2)dxdy$ et $\psi(t) = \iint_{D_t} \cos(x^2+y^2)dxdy$.

   Montrer que, quand $T \longrightarrow +\infty$, $\frac{1}{T}\int_0^T \phi(t)dt \longrightarrow \pi/4$ et $\frac{1}{T}\int_0^T \psi(t)dt \longrightarrow 0$

2. On choisit alors f afin que $D_1 = [0,1]^2$. Pour tout $t>0$, on pose

   $$C(t) = \int_0^t \cos(x^2)dx \text{ et } S(t) = \int_0^t \sin(x^2)dx$$

   Montrer que $\phi = 2SC$ et $\psi = C^2 - S^2$. En déduire la valeur des intégrales de Fresnel

   $$\int_0^{\to+\infty} \cos(x^2)dx \text{ et } \int_0^{\to+\infty} \sin(x^2)dx$$

EXERCICE 23-3.9 Chercher $f : \mathbb{R} \longrightarrow \mathbb{R}$ telle que la forme différentielle

$$\omega = \left(\frac{x^2 + y^2 - y}{xy} dx + \frac{dy}{y} \right) f(y/x)$$

soit exacte dans l'ouvert $\{(x, y) \in \mathbb{R}^2 \mid xy \neq 0\}$.

EXERCICE 23-3.10 On considère la courbe (C) d'équations paramétriques $x = \cos^5 t$ et $y = \sin^5 t$, ($t \in [-\pi, \pi]$). Calculer l'intégrale curviligne $\displaystyle\int_C \frac{x^3 dy - x^2 y dx}{(x^2 + y^2)^2}$.

EXERCICE 23-3.11 Construire la courbe plane d'équation : $(x^2 + y^2)^2 = xy$, puis calculer $\displaystyle\iint_D \sqrt{xy}\,dxdy$ où D représente le domaine intérieur à cette courbe.

Index

O

P

